1.0

1.1

1.25

1.4

1.6

1.8

2.0

2.2

2.5

2.8

3.2

3.6

4.0

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

AFOSR-TR- 81-0449

LEVEL

SIE/DIAL-81-005

(7)

OPTICAL COMPUTATIONS FOR IMAGE

BANDWIDTH COMPRESSION

by

B. R. Hunt

May 15, 1981

Final Technical Report

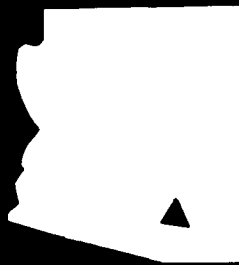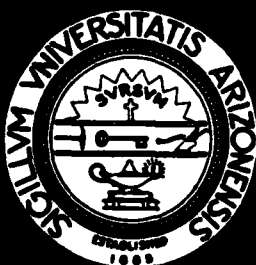Digital Image Analysis Laboratory

Department of Systems Engineering

University of Arizona

Tucson, Arizona 85721

ENGINEERING EXPERIMENT STATION

COLLEGE OF ENGINEERING

THE UNIVERSITY OF ARIZONA

TUCSON, ARIZONA

81 5 12 068

15 May 81

(I) REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM	
1. REPORT NUMBER AFOSR-TR-81-0449	2. GOVT ACCESSION NO. AD-A098823	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Optical Computations for Image Bandwidth Compression.		5. TYPE OF REPORT & PERIOD COVERED Final March 12, 1984 - March 12, 1984
		6. PERFORMING ORG. REPORT NUMBER SIE/DIAL-81-005
7. AUTHOR(s) B. R./Hunt		8. CONTRACT OR GRANT NUMBER(s) AFOSR-76-3024
9. PERFORMING ORGANIZATION NAME AND ADDRESS Digital Image Analysis Laboratory Department of Systems Engineering University of Arizona, Tucson, Az. 85721		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2305 B1
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NE Building 410 Bolling Air Force Base, D. C. 20332		12. REPORT DATE May 15, 1981
		13. NUMBER OF PAGES 87
14. MONITORING AGENCY NAME & ADDRESS(*if different from Controlling Office*) Final rept. 15 Mar 80-14 Mar. 81		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Optical Computations
Electro-Optical Computations
Image Bandwidth Compression
Spline/Polynomial Interpolation
Human Visual Models

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

This report summarizes current research in the investigation of optical computations for image bandwidth compression. The following topics are discussed:

(1)   A new adaptive image data compression scheme;

(2)   Simulations of an interframe adaptive image data compression scheme; and → next page

DD FORM 1473
1 JAN 73

1

Unclassified

(3) Simulations of an incoherent optical/
video feedback processor.

Table of Contents

(II) <u>Disclaimer</u>

    The findings in this report are solely those of the author and are not to be interpretted as the official position of the Air Force Office of Scientific Research or the U. S. Government.

(III) Introduction

Grant No. AFOSR-76-3024 was initiated on March 15, 1976. The overall goal of research sponsored under this Grant can be summarized in two different objectives:

(1) To demonstrate the feasibility of optical computations for the implementation of image bandwidth compression;

(2) Upon demonstration of feasibility, to carry-out the research necessary to advance optical computations for image bandwidth compression to approximately the same state of sophistication as currently associated with digital computations for image bandwidth compression.

Achieving these two objectives required the undertaking of several specific steps in the course of the research sponsored under the Grant. The steps were:

(1) Survey optical computation systems and schemes to establish the repertoire of functions which could be carried out optically, e.g., convolutions (both coherent and incoherent), sampling, differentiation, integration, etc.;

(2) Establish what computational processes or functions were useful in the development of bandwidth compression schemes;

(3) Determine architectural configurations for image bandwidth compression, that is, try to find

arrangements of optical processor components to

achieve compression calculations;

(4) Demonstrate, by means of simulations, that a
proposed optical processor architecture would
achieve some level of bandwidth compression;

(5) Seek improved architectures or new processor
functions to increase the compression perform-
ance of candidate systems.

These five steps were repeated continually for a variety of pro-
posed optical schemes for image bandwidth compression algorithms.
In this report, the final technical report for Grant AFOSR-76-3024,
we present the results of the efforts which were sponsored by the
Grant.

It is important to note that step (4) in the above list is
an important facet of our research, and is something which dis-
tinguishes our research from other research in both optical proc-
essing and image bandwidth compression. Every candidate archi-
tecture for bandwidth compression by optical processor was simu-
lated by using a multi-purpose, high-performance digital image
processing facility at the University of Arizona. (See Appendix A.)
It is important to understand the motivation for using digital
image processing in the development of optical processing schemes.

Optical processing research and development can be divided
into two broad categories: the development of physical materials,
devices, or systems which realize specific mathematical functions
or computational processes; and the development of systems archi-
tectures which treat the individual mathematical functions or

computational processes as components, and then assemble the components into a high-level system which achieves some complex processing objective.

Much of the research in optical processing falls into the first of these two categories: development of materials, devices, or systems which realize specific mathematical functions or processes. The concentration of effort in this area is quite important, for many of the great advantages of optical processing will not be realized without, for example, better materials to serve as spatial light modulators, or higher bandwidth in the schemes by which a modulator is accessed and addressed. The research in this area requires painstaking and careful control of experimental facilities.

In the research sponsored under Grant AFOSR-76-3024, we have concentrated in the second category: the development of systems architures. Because the devices and materials for optical processing are undergoing continual ferment, any attempt to physically realize an optical processor for bandwidth compression could be dependent upon the component devices chosen, e.g., modulator. In our research the use of digital image processing to simulate optical processor bandwidth compression systems is a direct consequence of our intent to concentrate upon systems-level concepts and avoid the turmoil of specific material or device implementation technologies. Digital image processing concentrates upon functional capabilities and their implementation, and is much more flexible. It is important to realize, however, that the use of digital image processing has been tied to optical realizability;

that is, no digital image processing functions were used in a simulation unless they could be realistically implemented in some properly configured optical device.

It is worthwhile noting that our course of investigation has gone from systems using all optical componentry to systems using a mix of optical and digital componentry. This is parallel to the path followed by other workers in optical processing. As the price, performance, and ease of interfacing of digital components and sensors has improved, more advantages to including digital components in optical systems have emerged. Hybrid optical/digital systems are being considered as the most effective solution to a number of signal-processing problems, and our work under Grant No. AFOSR-76-3024 is no different. The more complex systems for image gandwidth compression which have been studied in the latter phases of Grant AFOSR-76-3024 have explicitly involved the employment of digital componentry functions.

(IV) Summary of Important Results

We will summarize our important results first in terms of the general objectives set forth in the beginning of Section III, the Introduction:

(1) We have demonstrated the feasibility of a number of optical processing system architectures which can be employed for image bandwidth compression;

(2) Our results appear to justify the assertion that optical processing for image bandwidth compression can achieve the same sophistication as digital processing for image bandwidth compression.

We will summarize the specific systems which lead us to the results stated under item (1) directly above. First, we consider the result asserted in item (2), since this is a proper entry to describe the current state of digital processing for image bandwidth compression, and to describe the relationship of our research to it.

Image bandwidth compression schemes can be classified in two different ways. One major level of classification is in the domain where the compression scheme operates. There are two major divisions to this classification: space domain compression, where computations take place in the original image space of the input data; and transform domain compression, where computations are used to generate a transform of the original image, and subsequent compression computations take place in this transform domain, e.g., the Fourier domain. For either spatial or transform, the purpose of the compression computations is to remove the redundancy which

exists in the imagery, thereby producing an image which is "de-correlated" in some sense.

It is the dimension wherein the decorrelation takes place that defines the other major level of classification for bandwidth compression schemes. That is, the characteristics of the imagery are used to determine what decorrelation computations are carried out. For monochrome, single-frame imagery only spatial redundancy can be eliminated. For polychrome, single-frame images redundancy in both spatial and wavelength (or spectral) content can be eliminated. For multi-frame imagery, with each frame sequenced in time, any temporal redundancy can be eliminated. Combinations of any one or two of the redundancy measures of spatial, spectral, or temporal redundancy lead to different data compression schemes.

Digital image bandwidth compression schemes have been demonstrated for virtually all combinations of spatial or transform domain processing with spatial, spectral, and temporal redundancy reduction. It is this wealth of results which is the mainstay of literature in the compression of image bandwidth by digital computations.

A final sophistication in digital image bandwidth compression techniques is whether a given technique is adaptive or nonadaptive. A nonadaptive compression scheme processes all portions of an image in the same way, i.e., the operating parameters of the compression algorithm are the same for all components of the image. Conversely, an adaptive scheme recognizes that the level of redundancy of an image is not constant in space, spectral content, or time; for most

*See, for examples, the bandwidth compression chapters of Pratt, Digital Image Processing, Wiley, New York, 1978.

efficient operation, therefore, some of the critical operating parameters of the bandwidth compression algorithm are adjusted as a function of the behavior of the imagery.

Thus, digital image bandwidth compression processes can be characterized as: spatial domain vs. transform domain; spatial redundancy reduction vs. spectral redundancy reduction vs. temporal redundancy reduction; adaptive processing vs. nonadaptive processing. The overall complexity is graphically displayed in Figure 1. To add further to the complexity of digital image bandwidth compression schemes, recall that any of the schemes within a box in Figure 1 can be combined with others. Thus, a scheme could be developed, for example, by combining together nonadaptive spatial redundancy reduction processing in the spatial domain, with adaptive transform domain processing of the color or spectral components, and nonadaptive spatial domain processing to reduce temporal redundancy.

The point of Figure 1 is that it frames the challenge which is confronted in the second research objective in the introduction to Section III, i.e., to make the level of optical processing for image bandwidth compression equal to that of digital processing for image bandwidth compression.

The research conducted under Grant AFOSR-76-3024 has concentrated only upon the left-half of Figure 1, spatial domain processing. This is not because optics are not suited to transform domain processing. On the contrary, the Fourier transform capabilities of coherent optical systems are well suited to the computations for data compression which have been proven in the

	SPATIAL DOMAIN PROCESSING		TRANSFORM DOMAIN PROCESSING	
	NONADAPTIVE	ADAPTIVE	NONADAPTIVE	ADAPTIVE
Spatial Redundancy Only	Nonadaptive Intraframe Coding e.g., DPCM	Adaptive Intraframe Coding, e.g., Adaptive DPCM	Nonadaptive Intraframe Coding e.g., Cosine	Adaptive Intraframe Coding, e.g., Adaptive Cosine
Spectral Redundancy Only	Nonadaptive Multispectral Coding in Spatial Domain	Adaptive Multispectral Coding in Spatial Domain	Nonadaptive Multispectral Coding in Transform Domain, e.g., Cosine Spectral	Adaptive Multispectral Coding in Transform Domain e.g., Adaptive Cosine Spectral
Temporal Redundancy Only	Nonadaptive Interframe Coding, e.g., Temporal Frame DPCM	Adaptive Interframe Coding, e.g., Temporal Adaptive Frame DPCM	Nonadaptive Transform Temporal Interframe Coding e.g., Temporal Interframe Cosine	Adaptive Transform Temporal Interframe Coding, e.g., Temporal Adaptive Interframe Cosine

Figure 1

context of digital compression research. In fact, digital compression research using the transform domain is so extensive that it would be direct to develop a coherent system for carrying out the transform domain compression process. Such a system would have one severe problem: the requirement to employ coherent (holographic) detection in the Fourier plane in order to obtain both magnitude and phase components for the compression coding step. However, it could be done. Since this did not represent a challenge in the context of new and/or unique system-level architectures for optical processing in image bandwidth compression, such research was not undertaken. It is in the spatial domain processing, for spatial, spectral, and temporal redundancy, where our research was concentrated.

In the following we summarize the different optical processing schemes which were investigated during the course of research sponsored under Grant No. AFOSR-76-3024. The details of the actual processing schemes and the simulations are contained in the series of appendices affixed to this report. The appendix which documents a scheme summarized below should be consulted for specifics.

(IV.1) Interpolated DPCM

The simplest digital data compression scheme for single-frame monochrome imagery is Differential Pulse Code Modulation (DPCM). The first research undertaken for the purposes of Grant AFOSR-76-3024 was to construct a compression system which was the optical analogy of digital DPCM. In the process new insight was gained into digital DPCM as well.

The basis of digital DPCM is prediction. As the imagery is scanned (left-to-right, top-to-bottom), the pixels in a given causal neighborhood (called the prediction neighborhoood) are combined to form an estimate of the pixel which is next to be encountered in the scan. The difference between the estimate and the actual pixel value is then quantized and coded. DPCM works because, except in the neighborhood of abrupt slopes, an estimate will be accurate. Hence, the differential will be small and can be coded with a small number of bits.

The development of an optical analogy required releasing the causality constraint associated with an imagery scan, since optical systems are noncausal. In doing this, we replaced the pixel estimate with a neighborhood smoothing, and computation of a differential between the smoothed and actual pixel values. The differential proved to be small and was coded as in standard DPCM. Optical system configurations which would realize this new scheme, called interpolated DPCM, were configured on the basis of incoherent processors and video electronics.

Appendix B is a paper which described the IDPCM process in some detail.

(IV.2) Incoherent Feedback Video Processor

The noncausal optical system for IDPCM was then examined from the viewpoint of a generalization. It can be shown* that DPCM can be formulated as a temporal feedback scheme. A generalization of this scheme to imagery would require noncausal convolutions and

*See, for example, the book by Pratt, Digital Image Processing, Wiley, New York, 1978.

image-plane to image-plane feedback. An architecture to do this can be postulated as in Figure 2, where we show the general structure of causal DPCM and an architecture for noncausal image-plane to image-plane differences in an analogous fashion.

In theory the process in the lower half of Figure 2 can be realized with coherent processing, with beam splitters and phase adjustments serving as sums and differences; the quantization process could be realized by any of the nonlinear methods studied by Sawchuk*. However, the practical implementation of this process represents great difficulty, chiefly due to phase coherence. In even a miniaturized version of the scheme with integrated optics, the distance around the feedback paths will be very great compared with the wavelengths of light. The phase of signals at the difference and sum planes cannot be controlled, as a result, without resorting to interferometric precision.

Despite the difficulty of the feedback architecture, it was promising to investigate whether the phase problems could be solved. One direct choice is to employ video systems. By imaging onto video sensors and using the video signals for image-plane to image-plane operations the phase problem related to distance around the loop is solved, since video systems can operate at wavelengths which are long compared to the feedback path. A new problem is introduced, however, by the temporal scanning in a video system. Phase stability in spatial distance around the loop is exchanged for temporal stability caused by the time delay to scan one frame.

*Sawchuk and Dashiell, SPIE Proceedings on Image Processing, Vol. 74, p. 93, 1976.
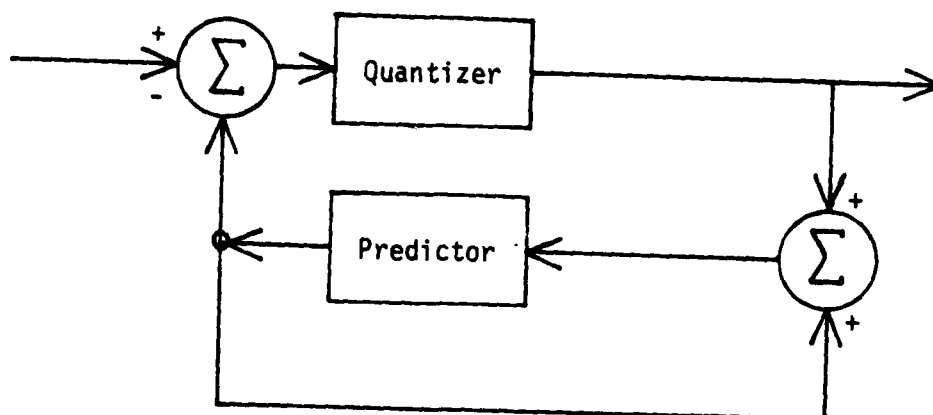
Figure 2-a.   Causal DPCM
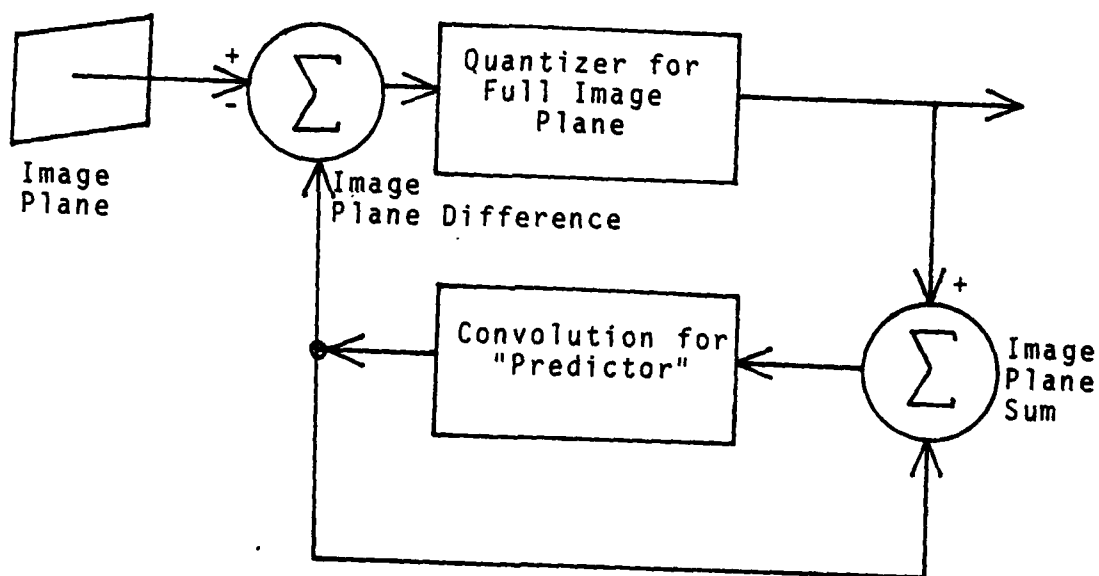
Figure 2-b.   Noncausal Image-Plane DPCM

Figure 2

We chose to investigate the problems of a sampled-data implementation of an incoherent optical video feedback processor. We also chose to simulate the operation of such a system. Our analysis is given in Appendix C. Our simulations showed the concept was feasible, but fraught with sufficient difficulties to make it unattractive. This line of research was discontinued.

(IV.3) Optical Image Bandwidth Compression in the Eye

The work which we conducted in the development and evaluation of IDPCM was further extended to an interesting insight into the human visual system. Based on a number of facts about the physiology of the human visual system, it is possible to recognize that the structure of the IDPCM bandwidth compression processor is quite similar to equivalent structure in the eye. Furthermore, both processes are suited for operation in incoherent light, which is an advantage for IDPCM and an absolute necessity for the human eye.

Once this analogy was recognized, research was undertaken to examine the extent to which the processing of information by the human visual system could be interpreted in a bandwidth compression context. Our results in this research are summarized in Appendix D, a paper discussing a two-channel model of human vision and its direct relationship to the IDPCM compression technique presented in Appendix B.

(IV.4) Spline Interpolation for Image Bandwidth Compression

With the success of the IDPCM scheme, which is nonadaptive and operates with incoherent light, the next step in complexity

was to introduce some degree of adaptive processing into the optical processors for bandwidth compression. Our research in this area was motivated by recent work in the adaptive compression of imagery by B-spline functions. The B-spline functions have a local property and can be constructed from repeated convolutions (which an optical processor can implement directly). Likewise, an analysis of the B-spline approximation problem shows that least-squares fitting of data can be derived by simple integration and differentiation processes, which an optical processor can also implement. Since any of these processes require coherent operation, however, the resulting processor must operate in coherent light and an incoherent-to-coherent conversion device (such as Hughes liquid crystal) becomes necessary.

Appendix E contains the detailed theory of an optical processor scheme which can adapt its compression behavior to the specific properties of the image in a localized region. It is a complicated processor, unfortunately. Although our simulations indicate that such a processor would have high performance in image bandwidth compression, we believe a viable adaptive processor must be more simple. This line of research was discontinued, as a result.

(IV.5) Adaptive Bandwidth Compression Scheme

Since the spline processor achieved adaptive compression behavior only at the expense of great complexity, a search for a much simpler scheme was undertaken. This led to the development of a scheme which utilized known behavior of the visual system:

sensitivity to edges. Many judgements about image quality, and many tasks in extraction of image information, are highly dependent upon the edge content of the image. An image can be rendered poor in contrast or resolution, but preservation of edges will still make the imagery suitable for many human purposes.

The adaptive scheme was developed on a requirement to preserve edges. A simple edge filter (which can be implemented optically) is created by convolution. The existence of an edge in the image is assumed if the edge strength exceeds a certain threshold. The position of the edges are coded by a run-length code, along with the edge values. A low-pass filtered version of the image (which can also be implemented optically) is transmitted with little bandwidth requirement. The edges are inserted into the low-pass version, and the result is an image with soft-shapes that possess sharp edges. The reconstructed images possess nearly all the information used in typical information extraction tasks- e.g., identification of objects.

Appendix F is a paper which describes the adaptive scheme resulting from this edge detection process.

(IV.6) Interframe Bandwidth Compression

Interframe imagery, i.e., imagery from a temporal, sequential image source such as television, represents a new level of complexity in the use of optical processors for image bandwidth compression. The problem is that temporal redundancy cannot be eliminated without storage of one or more frames prior to the current frame. It is only possible to identify temporal redundancy

in the context of the change from frame-to-frame.   .fortunately,
frame storage is difficult in an optical processor system. Al-
though electro-optical devices exist which can store a frame,
they are of inferior quality when compared to some of the digital
devices which have been recently developed, e.g., digital semi-
conductor frame buffer memories. Some new devices, such as CCD's,
offer analog sensing and digital addressing in simple, integrated
devices. Consequently, our research in interframe data compres-
sion has been on the basis of hybrid optical/digital operation.
The employment of digital frame buffers to serve as image storage
is the simplest way to allow for the temporal processing align-
ment.

Appendix G is a paper describing our system proposed for in-
terframe compression. Basically, it consists of an optical com-
pression component utilizing the IDPCM process discussed in Ap-
pendix B. Since multiple frame imagery is assumed, the output of
the IDPCM processor is captured in frame buffer units, and con-
ventional DPCM processing between frames is used to reduce the
redundancy between frames.

(V) Conclusions

We belive the papers and results set forth in detail in the Appendices justify the conclusion that we have demonstrated the feasibility of using optical processors for image bandwidth compression. This was the first objective which we set forth in the introduction, Section II. What of our second objective, to raise the sophistication of optical processing methods to the level enjoyed by digital bandwith compression? As can be seen from Figure 1 of Section IV, we need additional research to establish a breadth of variety in optical computations for image bandwidth compression, comparable to digital systems. In particular, bandwidth compression for multi-spectral imagery, better interframe compression performance, and adaptive compression are all goals for our future research in this area.

(VI) <u>Publications</u>

The following publications resulted from either whole or partial support under Grant No. AFOSR-76-3024.

(1) Hunt, B. R., "Some remaining mathematical problems in nonlinear image restoration", <u>Proceedings Image Science Symposium</u>, Naval Post-Graduate School, Monterey, Western Publishing Co., 1976.

(2) Granrath, D. J. and Hunt, B. R., "Signal-detection trade-off-analysis of optical vs. digital Fourier transform computers", <u>Applied Optics</u>, Vol. 18, pp. 36-43, 1979.

(3) Hunt, B.R., "Optical computing for image bandwidth compression: analysis and simulation", <u>Applied Optics</u>, Vol. 17, pp. 2944-2951, 1978.

(4) Granrath, D. J. and Hunt, B. R., "A two-channel model of image processing in the human retina", <u>SPIE Proceedings</u>, Vol. 199, San Diego, August, 1979.

(5) McCaughey, D. G., "An image coding algorithm using spline functions", <u>SPIE Proceedings</u>, Vol. 149, San Diego, August, 1978.

(6) Hunt, B. R., "Non-stationery image models (and their application to image data compression)", <u>Computer Graphics and Image Processing</u>, Vol. 12, pp. 173-186, 1980.

(7) Cabrera, S. D. and Hunt, B. R., "Optical implementation of a spatially adaptive image data compression system", to be published, <u>Optical Engineering</u>, 1981.

(8) McCaughey, D. G. and Andrews, H. C., "Image approximation by variable knot bicubic splines", to be published, <u>IEEE Trans. Computers</u>, 1981.

(9) McCaughey, D. G. and Andrews, H. C., "The continuous-discrete model: least-squares inverses and singular function expansions", to be submitted for publication, <u>IEEE Trans. Info. Theory</u>.

(10) McCaughey, D. G. and Hunt, B. R., "Spatially adaptive optical data compression by the finite element method", in preparation.

(VII) <u>Personnel</u>

The following students were supported under the auspices of Grant No. AFOSR-76-3024:

(1) Everitt Sherwood, M.Sc. Degree, 1976.

(2) Paul McGooey, M.Sc. Degree, 1977.

(3) Douglas Granrath, Ph.D Degree, 1979.

(4) S. D. Cabrera, M.Sc. Degree, 1979.

(5) John Tressler, Ph.D Degree, 1981 (expected).

(6) Hiroyasu Ito, Ph.D student, 1979-Present.

(7) Richard Stevens, Ph.D Candidate, 1980-Present.

The following faculty received some support from Grant No. AFOSR-76-3024:

(1) B. R. Hunt, Principal Investigator, Professor of Engineering and Optical Sciences (supported from 1976 to 1981).

(2) D. G. McCaughey, Assistant Professor of Engineering (support during 1977-78).

(3) R. L. Baker, Associate Professor of Engineering, (support during 1977-78).

APPENDIX A

THE DIGITAL IMAGE ANALYSIS LABORATORY

THE UNIVERSITY OF ARIZONA

T U C S O N , A R I Z O N A 85721

SYSTEMS ENGINEERING DEPARTMENT

THE DIGITAL IMAGE ANALYSIS LABORATORY

The Digital Image Analysis Laboratory (DIAL) is located in the Engineering Building on the University of Arizona. campus. The DIAL facility is the focus on the Arizona campus of research involving the processing, manipulation, and analysis of imagery by digital computers. Basic research is also carried-out in technologies that support digital image analysis, such as signal processing and numerical techniques.

Besides the use of DIAL facilities in programs of sponsored research, DIAL figures prominently in the education of graduate students. DIAL facilities are used for instruction in Remote Sensing, Optical Sciences, and Systems Engineering. Students from these Departments (as well as others) regularly work in the DIAL facility.

DIAL Resources

The resources within DIAL consist of equipment, programs, and faculty of the University of Arizona.

(1) Equipment

DIAL equipment resources consist of hardware physically located within the laboratory, and of hardware outside of the laboratory that can be accessed remotely. Equipment with DIAL consists of the following:

- PDP-11/70 Computer with 192K words of storage, cache memory, floating point processor.

- RM03 Disc Storage Unit, with 67 Megabytes of memory.

- TU-10 Magnetic Tape Drive, 9 track 800 BPI.

- DZ-11 8-line Multiplexer, with 4 CRT terminals, one printing terminal, one remote dial-up phone line modem terminal.

- LP-11 Line Printer.

- Stanford Technology Corporation Image Display Unit: 512 x 512 pixels x 8 bits/pixel x 3 colors (Red, Green, Blue), with Graphics, Feedback Arith./Logic Unit, Precision color CRT display monitor, interactive trackball control.

Equipment accessed remotely from DIAL includes the following:

- DEC-10/CDC Cyber 175 Computers in the University Computer Center; access from DIAL is a permanent 1200 baud telephone terminal modem set.

- VAX-11/780 Computer in the Department of Radiology; access is a 19.2 k baud telephone modem set between the VAX and 11/70 CPU's.

(2) Software/Programs

DIAL software resources consist of special programs written by equipment manufacturers and programs created for more general-purpose image processing. Software resources include:
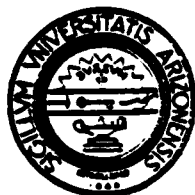
SYSTEMS ENGINEERING DEPARTMENT

- SADIE, a general-purpose image processing software package written at the University of Arizona and distributed by the University of Minnesota to a dozen different sites. SADIE is available in both a CDC Cyber 175 version and a PDP-11/70 version;

- FACEL, a general purpose pattern recognition package;

- System 511, an interactive image processing package solely for the Stanford Technology Corporation image display system.

(3) Faculty

Faculty working in DIAL, or associated with DIAL projects, cover a variety of academic disciplines. Principals currently involved include:

B. R. Hunt, Professor of Systems Engineering and Professor of Optical Sciences;

P. N. Slater, Professor of Optical Sciences and Chairman of Remote Sensing Programs;

J. J. Burke, Professor of Optical Sciences;

R. Schowengerdt, Assistant Professor of Remote Sensing.

Some DIAL Activities

A variety of sponsored research projects are currently underway within DIAL. Examples of the current projects are:

- Simulation of optical processing for image data compression;

- Investigation of factors influencing the automatic compilation of maps from aerial photographs;

- Study Group to examine the feasibility of automating image processing;

- Geometric correction, rectification, and editing of images of the planet Saturn from the Pioneer spacecraft;

- Generation and display of imagery for ranking of image quality criteria;

- Use of LANDSAT imagery for determination of agricultural planting patterns in Avra Valley, Arizona.

Further Inquiries

For further information, inquiries concerning research projects, sponsorship, or use of DIAL facilities, please contact:

Professor B. R. Hunt
DIAL
Department of Systems Engineering
University of Arizona
Tucson, Arizona 85721
(602)626-5157

APPENDIX B

OPTICAL COMPUTING FOR IMAGE BANDWIDTH

COMPRESSION: ANALYSIS AND SIMULATION

Optical computing for image bandwidth compression: analysis and simulation

B. R. Hunt

Image bandwidth compression is dominated by digital methods for carrying out the required computations. This paper discusses the general problem of using optics to realize the computations in bandwidth compression. A common method of digital bandwidth compression, feedback differential pulse code modulation (DPCM), is reviewed, and the obstacles to making a direct optical analogy to feedback DPCM are discussed. Instead of a direct optical analogy to DPCM, an optical system which captures the essential features of DPCM without optical feedback is introduced. The essential features of this incoherent optical system are encoding of low-frequency information and generation of difference samples which can be coded with a small number of bits. A simulation of this optical system by means of digital image processing is presented, and performance data are also included.

I. Introduction

When digital image processing methods were initially employed ten to fifteen years ago, the rationale often voiced in making the decision to use digital computations was the flexibility of the computer and the relative inflexibility of optics. Thus, it was often argued, the digital computer offered the means to explore and simulate a variety of system configurations. Once the final optimum configuration became known, this particular one could be frozen and implemented in optics, which have the virtues of parallel computations and wide space–bandwidth product. This early rationale for digital image processing is heard with much less frequency nowadays. The revolution in semiconductor electronics has produced cheap, fast, and reliable digital systems. New digital algorithms, such as the fast Fourier transform, have made digital image processing an active and fruitful endeavor, which is carried on for its own purposes and without reference to flexibility in simulating optics. In this recent burst of activity in digital image processing, a remembrance of the early rationale brings to mind the question: are there optical computations in image processing which are being overlooked in the successes of digital processing?

Image data compression is an example of the success

The author is with University of Arizona, Systems & Industrial Engineering Department and Optical Sciences Center, Tucson, Arizona 85721.

of digital image processing. The motivation for data compression is the great amount of information that can exist in an image. Even a low-quality image, such as might be produced by a pocket camera, can contain 10^7–10^8 bits of information, and a high quality image can contain several orders of magnitude more bits of information. The transmission and storage of such masses of data are difficult, and anything which can be done to eliminate data redundancy is of interest, since there will be a concurrent reduction in the requirements of transmission bandwidth, storage, and system costs. Given this motivation, image data compression has been one of the most successful applications of digital image processing. As can be seen from existing survey papers on image data compression,[1-3] a variety of different methods have been investigated and shown to be successful. The success of digital computations can be seen in efforts currently underway to build and test prototypes of all-digital compression systems for military applications.[4,5]

As favorable as the performance characteristics of current digital components are, the potential of optical computations should not be overlooked. There is still merit in analog signal processing when the task is properly defined. In this paper we discuss the application of optical methods to the problem of image data compression. We will consider the computations employed in image data compression from the viewpoint of how those computations may be realized by optical processes. In addition, a structure suitable for optical implementation of image data compression will be presented. The results of simulating this proposed optical system by digital image processing will be presented. (This brings things to full circle, i.e., simulating
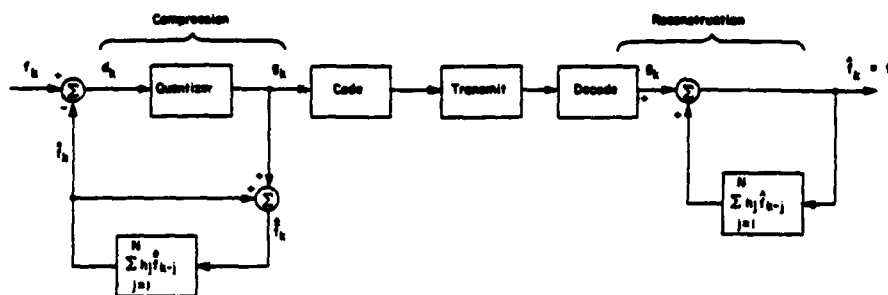
Fig. 1. Schematic of DPCM digital compression.

possible optical systems by digital processes!) Thus, this paper has two aims: to discuss the applicability of optical computations in image data compression and to present results of simulating such computations.

II. DPCM Digital Image Data Compression

In approaching the problem of performing image data compression by optical computations, we begin by a brief review of the principles behind the simplest of image compression schemes, differential pulse code modulation (DPCM). This review is necessary to motivate the optical analogies we present below.

The conventional structure of a DPCM image data compression system is seen in Fig. 1. The data compression process takes place in the left-hand feedback loop, and reconstruction occurs in the loop at the right. The basic equations are the following. The past N samples of previously predicted data are used in a linear predictor to generate a prediction of the current sample

$$\hat{f}_k = \sum_{j=1}^{N} h_j \hat{f}_{k-j},$$ (1)

where k is the index of the current sample, h_j is the prediction weight given to each of the previous predicted samples, and $\hat{f}_k$ is the prediction of the value of the kth sample. (From Fig. 1, we see that $\hat{f}_k = \hat{f}_k + g_k$, which is the relation between a predicted pixel and the quantized value of a predicted pixel.) The difference between the kth sample and the predicted value of the kth sample is computed

$$d_k = f_k - \hat{f}_k,$$ (2)

and this difference is quantized

$$g_k = Q(d_k),$$ (3)

where Q is the quantizer function. The quantized difference g_k is then coded for transmission as well as being fed back around the quantizer, where it becomes input to the prediction computations. As Fig. 1 shows, the reconstruction process consists of a positive feedback and combination of the decoded differences with the output of a prediction computation which is identical to that in the original compression loop.

The workings of DPCM are extremely simple but are sometimes difficult to understand because of the feedback occurring around the quantizer. To show the basic workings of DPCM we resort to a z transform analysis. First we replace the quantizer in the compression system by the injection of an additive source of quantization noise n_k as seen in Fig. 2. That is, the quantizer is modeled as a linear addition of noise n_k to the kth input of the quantizer. The replacement of a quantizer by an additive noise source is a standard assumption in digital signal processing.[6] We use the z transform notation that

$$f_k \xleftrightarrow{z} F(z)$$ (4)

and likewise for other symbols. In Fig. 2 we have two feedback loops, one at node 1 and one at node 2. For node 1 it is direct to write the relation

$$\hat{f}_k = \sum_{j=1}^{N} h_j(g_{k-j} + \hat{f}_{k-j}).$$ (5)

Using the z transform property of discrete convolutions[6] (identical to that for Fourier transforms and continuous convolutions), we have

$$\hat{F}(z) = H(z)[G(z) + \hat{F}(z)].$$ (6)

Solving for $\hat{F}(z)$ the result is

$$\hat{F}(z) = \frac{H(z)G(z)}{1 - H(z)}.$$ (7)

Likewise at node 2 we have (in z transforms)

$$G(z) = F(z) - \hat{F}(z) + N(z) = F(z) - \frac{H(z)G(z)}{1 - H(z)} + N(z).$$ (8)

Solving for $G(z)$ we have the result
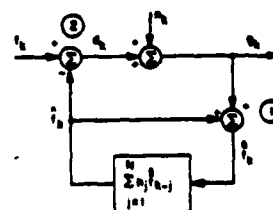
$$G(z) = [1 - H(z)][F(z) + N(z)].$$ (9)

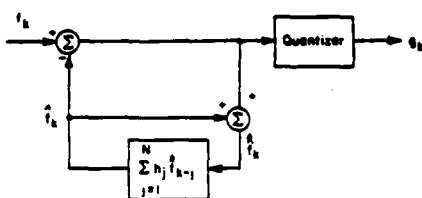Fig. 2. Replacement of quantizer by additive noise.

Fig. 3. Alternate configuration with quantizer outside of feedback loop.

The reconstruction system at the receiver is directly seen to have the z transform description (assuming there are no transmission or decoding errors)

$$\hat{F}(z) = \frac{G(z)}{1 - H(z)},\qquad(10)$$

and the resulting output of the over-all system is

$$F_r(z) = F(z) + N(z).\qquad(11)$$

We see that the total compression system adds only quantization noise to the reconstructed data. We also see the importance of the structure of the compression loop. If the quantizer were outside the feedback loop as in Fig. 3, the equivalent of Eq. (9) would be

$$G(z) = [1 - H(z)]F(z) + N(z),\qquad(12)$$

and the reconstructed data would have the z transform

$$\hat{F}(z) = F(z) + \frac{N(z)}{1 - H(z)}.\qquad(13)$$

The differences between Eqs. (11) and (13) can be used to highlight the basic operations of DPCM. The optimal design of the predictor weights h_j can be carried out by minimizing the mean-square-error of prediction.[2] Since it is possible to model images as Markov processes of small finite order (e.g., third order), the optimal weights for h_j usually result in a low-pass transfer function.[7] The transfer function $1 - H(z)$ is thus a high pass transfer function. This is expected behavior, since an accurate prediction $\hat{f}_k$ of the current sample f_k implies that the difference d_k will consist only of features that cannot be predicted, i.e., high frequency image details that are not predictable from the general low-frequency trends in the image data. Further, since $[1 - H(z)]$ is a high pass process, the quotient $1/[1 - H(z)]$ will have appreciable amplitude at low frequencies, decreasing at higher frequencies, i.e., it is the frequency characteristic of an integrating filter. In Eq. (11), the quantizer being within the feedback loop causes the transfer function of the integrating filter to be completely canceled. In Eq. (13), however, the integrating filter is actually integrating the quantization noise throughout the image data, and the actual reconstruction errors are much greater.

The salient points of DPCM compression can be summarized in the following three properties:

(1) The prediction and differencing steps remove

those portions of the image which are predictable on the basis of past samples, i.e., the low-frequency information. This low-frequency information is recreated in the reconstruction process by the integrating filter.

(2) If the prediction is accurate, the difference d_k will be much smaller in absolute magnitude than the incoming data samples f_k. Since the quantization noise is directly proportional to the variance of the input,[8] the error induced by quantization of the smaller amplitude differences is decreased. For example, if the original samples f_k spanned a range of 1–1000 (on some arbitrary scale), quantization accuracy of one part in 1000 would require 10 bits of quantization. However, if the prediction/difference process yields values d_k with a range of 1–10, the same absolute quantization accuracy could be obtained with 4-bit quantization and with appreciable reduction in code bits.

(3) The DPCM process is causal, i.e., the nature of prediction implies an ordering to the samples. This ordering is shown as 1-D in the equations and diagrams above, but a 2-D image is naturally ordered by a raster scan process such as used in conventional video systems. In a raster scan the 2-D data available to the prediction process are characterized as data to-the-left-and-above the current position of the scanning spot (assuming a conventional top-to-bottom left-to-right scanning system).

III. Direct Optical Analogies to Digital DPCM

Given the three properties of digital DPCM summarized above, it is immediately evident that one property of the digital process is not applicable to the optical case: the property of causality and the related processes of prediction. Causality is inherent in the prediction process, i.e., the past samples are used to compute a prediction of the next sample, a process which assumes that the next sample is unknown until the elapse of the next sample time. The close tie between this process and the raster scanning of an image is obvious.

Optical image formation occurs simultaneously over the image plane, and raster scanning of the image is artificially a causal ordering upon the totally parallel or noncausal process of image formation. A naive approach to constructing an optical analogy to DPCM
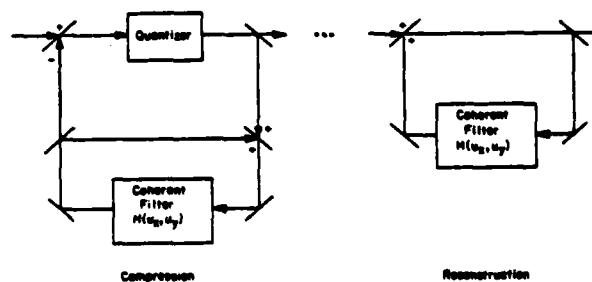

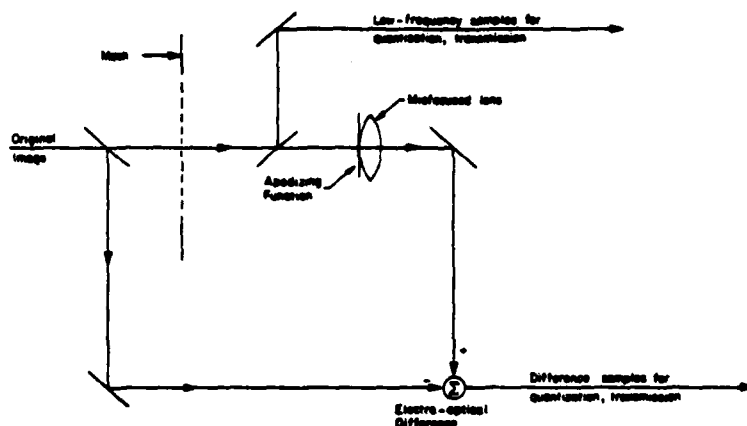Fig. 4. Coherent optical analogy to DPCM compression.

Fig. 5. Incoherent optical system of IDPCM compression.

would be to replace the causal system presented in Fig. 1 by a noncausal system with the same structure, i.e., a system having the same structure by functioning in the parallel optical mode. Such a hypothetical processor can be seen in Fig. 4. It is a coherent optical system with feedback. Fully parallel combination of images serve to create negative or positive feedback. In addition, a fully parallel optical quantizer is included within the feedback loop.

The equations which describe such a system can be written from the diagram of Fig. 4. It is direct to show that the equilibrium output (i.e., the output after an initial light wave passes around the system, and it settles) is given by

$$g(x,y) = Q\left[f(x,y) - \int_{-\infty}^{\infty}\int p(x - x_1, y - y_1) g(x_1,y_1) dx_1 dy_1 \right],$$

(14)

where $p(x,y)$ is the spread function of the filter in the lower loop, i.e.,

$$p(x,y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty}\int \frac{\exp[-i2\pi(u_x x + u_y y)]}{1 - H(u_x,u_y)} du_x du_y,$$ (15)

and Q is the quantizer function. Equations (14) and (15) describe the space-domain output of the noncausal compression system, i.e., the image plane emerging from the compression system is described in the space domain by the solution of the nonlinear integral Eq. (14), given the spread function defined in Eq. (15).

What can be stated about the feasibility of a system such as seen in Fig. 4? The following points are worth noting:

First, the coherent filters in both compression and reconstruction systems represent the best defined technology, and methods to achieve the filter response are described in detail in the literature.

Second, the fully parallel optical quantizer is an emerging technology. The nonlinear optical methods of Dashiell and Sawchuk[9] would be applicable to the synthesis of the quantizer, but the utilization of such a nonlinear element integrated into a system with other optics, including feedback, raises difficult questions, i.e.,

errors in the optical quantization process and cumulative effects of these errors after many trips around the loop, when the feedback system settles into its equilibrium state.

Third, optical feedback has been found to be possible only by painstaking measures to ensure component rigidity and stability, freedom from random perturbations in the optical path, control of the coherence length of the illumination, etc.[10,11] The problems arise because the system shown in Fig. 4 is, in essence, an image forming interferometer. The problems associated with it are not necessarily insurmountable, but they are not to be taken lightly.

IV. Incoherent Optical Compression System

At the close of Sec. II, three important points of digital DPCM were stated. The third point, causality, was dealt with in the conceptual formulation of Fig. 4, i.e., the causality constraint of digital DPCM was eliminated by a noncausal optical system having a schematic block diagram equivalent to digital DPCM. We now reexamine the other two points. DPCM works because (1) low frequencies in the image can be predicted, removed from transmission by a difference step, and then recreated at the receiver; and (2) the high frequencies, which are not eliminated by prediction and differencing, possess only a small amount of the total image energy and can be accurately coded with a small number of bits. Given these two points, the development of optical analogies to digital DPCM data compression can be perceived in terms of the coding of low-frequency information, and the coding of the differences between low-frequency and high frequency information.

There are many ways to encode the low-frequency information. A key feature is to encode the low-frequency information in such a way that low-frequency image details can be reconstructed at the receiver without undue amplification of the quantization noise. It is this requirement which makes necessary placing the quantizer inside the feedback loop in Fig. 4 and makes analyzing the system subsequently difficult. [Recall the discussion associated with Eq. (13) to see that

placing the quantizer outside of the feedback loop would amplify the quantization noise.] We choose to encode the low-frequency information by sampling and interpolation, since these processes are simple to implement optically.

Figure 5 shows a simple incoherent optical system for the process of image data compression. An image of the scene is formed on a mask, which is an opaque screen with small holes periodically spaced in it. The mask functions as an optical sampling element. The samples extracted from the image by the mask are then optically interpolated to create a low-frequency version of the original image. The interpolation is accomplished in a classical incoherent convolution[12] using an apodized aperture and a misfocused lens. Then the low-frequency version created by interpolation is subtracted from the original unsampled in-focus image. Finally, the samples from the mask and the difference image are quantized and transmitted.

The following points can be made about the detailed workings of the system:

(1) The samples are a crude representation of the original image, since the spacing of the samples would be chosen not to satisfy the Nyquist criterion but by a desire to represent low-frequency information in a small number of samples. For example, suppose the original image had a Nyquist imposed resolution of 512×512 pixels. Choosing a 128×128 sampling mask would encode the low frequencies at only $\frac{1}{16}$ the data requirements of the original image. Further data reductions are possible because these low-frequency samples need not be quantized at full resolution, i.e., instead of 8 bits/pixel, 3 or 4 bits/pixel could be used.

(2) The coarse sampling from the mask results in a sampled image which is badly aliased. Likewise, coarse quantization of the optical intensities sampled by the mask induces quantization error. However, both aliasing and quantization errors are encoded in the difference image (along with image high frequency information) and are reintroduced into the reconstructed image (see point 5).

(3) The optical interpolation acts upon that portion of the image least affected by the aliasing from the sampling—the low frequencies. A misfocused lens is suitable for image convolution by a low-pass filter.[12] As discussed in the previous sections, the general characteristic of a prediction filter in DPCM is a low-pass characteristic. In the system of Fig. 5 we replace the prediction filter of DPCM by an interpolation filter, i.e., the misfocused lens with associated aperture apodization interpolates between the samples to fill-in the gaps with a low-frequency version of the original image.

(4) The nature of the interpolated low-frequency version depends upon the interpolation criterion adopted. An interpolation criterion can be stochastic, e.g., minimization of mean-square-error can be used to derive a well known result, which determines the interpolator characteristics in terms of the autocorrelation function of the data.[13] A deterministic interpolation criterion can be specified to represent exactly a polynomial of nth order, and an attractive class of polynomial interpolating functions for images are the B splines, which can be implemented as convolutions.[14] Of course, if the image data possess an autocorrelation function which is identical to an nth order polynomial, these two different approaches (stochastic vs deterministic interpolation criteria) become equivalent.

(5) At the receiver the reconstruction is similar to the compression. Low-frequency samples are written onto an optical modulator, and identical interpolation is performed to create the low-frequency image. The low-frequency image is summed with the difference image to recreate an approximation to the original.

(6) All the computations to carry on the compression and reconstruction process can be carried on optically or electronically. The difference between the interpolated image and the original image can be performed either electronically (e.g., use two vidicons looking at the two images, with the synchronization signals slaved together, and create the difference by an operational amplifier) or can be performed electrooptically (e.g., using such electrooptic devices as the PROM or the liquid crystal[15,16]). The only digital circuitry required would be in the quantization of samples from the mask and from the difference image. Note that, since these will be coarse quantizations, 2–4 bits/sample, the associated A–D converters will be faster in performance than A–D converters used to quantize an analog scene at 8–10 bits for input to digital DPCM processing. The optical system of Fig. 5 should be able to operate at greater data rates and be significantly simpler than an equivalent digital DPCM device.

This method of creating a DPCM image is referred to as interpolated DPCM or IDPCM to distinguish it from ordinary DPCM. We summarize the entire process in the following:

First, the image is sampled by a mask, and the samples are quantized and transmitted.

Second, the sampled image is interpolated by a low-pass convolution

$$f_i(x,y) = h(x,y) * f_s(x,y), \qquad (16)$$

where h is the PSF of the interpolator, and f_s is the sampled image created in the first step.

Third, the difference image is computed

$$d(x,y) = f(x,y) - f_i(x,y), \qquad (17)$$

and all samples of it are quantized and transmitted.

Fourth, at the receiver the mask samples, i.e., the samples of f_s, are again interpolated to form a low-frequency version.

Fifth, the difference samples are added to the interpolator output to reconstruct the original.

A simple theory of the operation of the IDPCM system in Fig. 5 can be readily derived. After the mask samples have been quantized, we model the quantizer as in Sec. II and have

$$Q[f_s(x,y)] = f_s(x,y) + n_1(x,y), \qquad (18)$$

where n_1 is the noise from quantization of the mask samples. Likewise, a quantization acts upon the difference between the original image and the interpolated
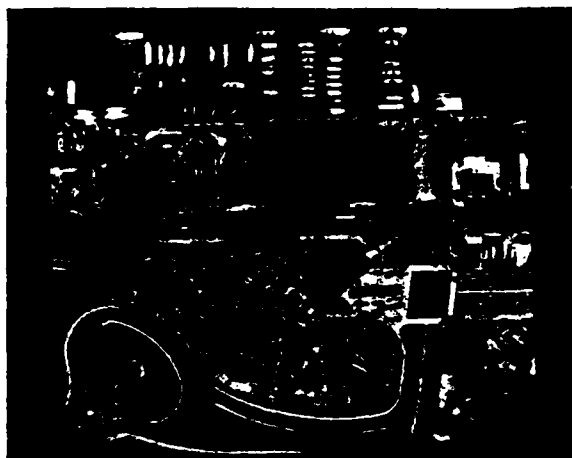
Fig. 6. Original image used in simulation of IDPCM compression.

mask samples, so this second quantization process is modeled as

$$Q[d(x,y)] = f(x,y) - h(x,y)*f_s(x,y) + n_2(x,y), \qquad (19)$$

where n_2 is the noise from the second quantization process. At the receiver, the received quantized mask samples are reinterpolated and added to the received quantized differences. Thus, the reconstructed image is

$$f_r(x,y) = h(x,y)*[f_s(x,y) + n_1(x,y)] + f(x,y),$$
$$- h(x,y)*f_s(x,y) + n_2(x,y),$$
$$= f(x,y) + h(x,y)*n_1(x,y) + n_2(x,y). \qquad (20)$$

This equation demonstrates that the reconstructed image has two noise contributions. The noise term n_2 is due to quantization of differences and is, hence, identical in nature to the quantized difference noise seen in Eq. (11). The other noise term is unique to the IDPCM process and is the convolution of the interpolating function with the noise introduced by quantization of the mask samples. The noise from a quantizer tends to have a zero mean (positive and negative values both occur[6]). The averaging of values which takes place in a low-pass convolution with the interpolating function results in the spatial average of the mask samples tending toward zero, particularly near sharp edges in the image. An area of the image where the underlying image structure is predominantly low frequency can give rise to quantization errors in that spatial region with a single algebraic sign or a predominance of one sign. In such cases the convolution with the interpolating function does not average out the quantization noise of mask samples in the image. The result is a visually noticeable error in representation of low-frequency image structure (e.g., regions of constant or near-constant intensity) but few visually detectable errors around sharp edges of the image. This latter fact is fortunate, since image edge information is usually the most sensitive in subjective viewer evaluations, as can

be seen in the images presented in the following section.

The important feature of this simple analysis of the IDPCM process is that it demonstrates that the reconstruction process does not grossly amplify the quantization noise present in the data, a severe problem with the nonfeedback quantization of Eq. (13). Indeed, the quantization noise in the mask samples is reduced by the convolution with the interpolating function, the actual magnitude of the reduction being dependent upon the nature of the interpolating function chosen and the specific region of the image.

V. Simulations of the Optical Compression Scheme

A series of digital image simulations have been carried out to verify the validity of the IDPCM process. Figure 6 is an original image, sampled at 9 bits/pixel on a 480 × 480 raster. The compression steps carried out in the digital simulation process were in the same sequence as described above. The original image was subsampled, retaining every 4th pixel of every 4th line and creating a 120 × 120 image of mask samples. Each subsample in this 120 × 120 array was quantized to 3 bits (8 levels) with a uniform quantizer, the maximum and minimum quantization levels being the maximum and minimum of the subsamples. Thus, the low frequencies were encoded with 120 × 120 × 3 bits of total information. These samples were transmitted.

The image intensities from the sample mask were interpolated to fill-in the missing data values. This was done by inserting zeros into the positions where pixels were missing and then convolving the resulting 480 × 480 array with a 7 × 7 bilinear interpolation kernel.[17] The interpolated image, being of the same 480 × 480 resolution as the original image, was now subtracted from the original, and the differences were quantized at N_D bits, which was varied. The quantization rule was a tapered quantizer, based upon the Laplacian density used by O'Neal in quantizing differences in digital DPCM.[7] The low-frequency mask samples and the quantized differences constituted the information that would be transmitted in a real system. The image was reconstructed, in a simulation of the receiver, by reinterpolating the low-frequency mask samples from 120 × 120 to 480 × 480 (using the same bilinear interpolator as in the transmitter simulation), and the result was added to the quantized differences to reconstruct the image.

Total bits required in the simulations can be calculated from the equation

$$B_{total} = \frac{480 \times 480}{16} \times 3 + 480 \times 480 \times N_D,$$
$$= \frac{3}{16} N^2 + N^2 N_D, \qquad (21)$$

where the original image is of size N by N, and N_D is the number of bits chosen to encode the differences. The number of bits/pixel is

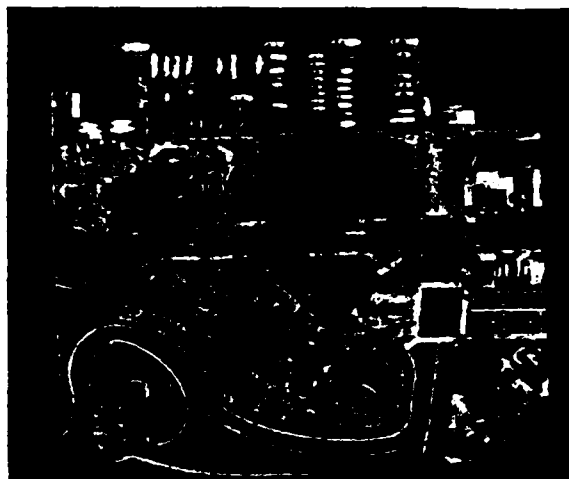$$B/pixel = \frac{3}{16} + N_D. \qquad (22)$$

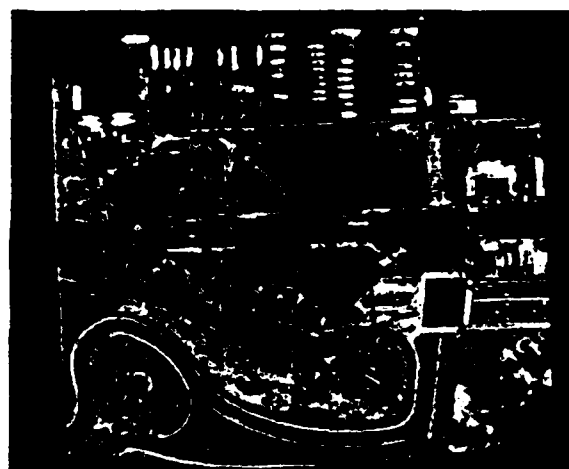Fig. 7. IDPCM compression ≈3 bits/pixel (see text).

Fig. 8. IDPCM compression ≈2 bits/pixel (see text).

In general, if N_L bits are used to encode the low-frequency mask samples and if the mask retains every Rth pixel of every Rth line, the above equations become

$$B_{total} = \frac{N^2}{R^2} N_L + N^2 N_D, \qquad (23)$$

$$B/pixel = \frac{N_L}{R^2} + N_D. \qquad (24)$$

The quantities N and R are interpretable in terms of the simulation discussed above. However, even in a real optical implementation of IDPCM, in which there is no image originally sampled at N by N resolution, the parameters N and R are still meaningful. Even though an optical system would not initially sample the image to be coded, it is still possible to express the information content of the original image in terms of an N by N sampling. For example, N by N could be the size of the

sample array of an image sampled at the Nyquist rate corresponding to the incoherent diffraction-limited frequency cutoff of the optical aperture. A choice of N by N less than the sampling rate at the diffraction limit would indicate a decision to allow a given degree of aliasing. In this context, the parameter N does not represent the sampling of an image such as used in the simulations being described but represents the information requirements of an image acquisition/transmission system with no data compression. Likewise, in the context of a real optical implementation of IDPCM, the parameter R represents the subsampling of the nominal Rth pixel of every Rth line for the data compression processes of IDPCM.

Figure 7 is the result of encoding and reconstruction for $N_D = 3$ bits/difference sample. The image is virtually identical to the original image. Figure 8 shows the same result for $N_D = 2$ bits/difference sample. Some distortion is visible in Fig. 8, but the over-all quality remains high. Figure 7 has normalized mean-square error of 0.3%. For Fig. 8 the error is 1.4%. The normalized mean-square-error is expressed as

$$NMSE = \frac{\sum_{x,y} [f_r(x,y) - f(x,y)]^2}{\sum_{x,y} f(x,y)^2}. \qquad (25)$$

The NMSE measure is commonly used to provide an error measure that is independent of the mean intensity level of the original image. It is worth noting that the NMSE values for the IDPCM process in Table I are comparable with NMSE values in conventional digital DPCM compression schemes.[1]

Table I summarizes the performance of the IDPCM simulation on Fig. 6. The compression ratio is the bits/pixel of the original image divided by the bits/pixel of the compressed image. Table I also tabulates the coding efficiency that could be achieved if a more sophisticated method of coding pixel differences is used. Pixel differences were found to be described accurately by a Laplacian probability density such as used by O'Neal in conventional DPCM.[7] Using 8-level and 4-level Huffman codes, based on the Laplacian density, gives the coding performance shown in Table I. A Huffman code saves an additional 0.5 bit (approximately), but the cost is in much greater complexity in dealing with the resulting variable length code words.

VI. Concluding Comments

The simulations presented above indicate the validity of relatively simple optical systems which can carry out

	Coded bpp	Error (%)	Compression ratio	Huffman code bpp	Huffman compression ratio
Fig. 7	3.19	0.3	2.82	2.48	3.63
Fig. 8	2.19	1.4	4.11	1.78	5.06

Table I. IDPCM Performance Summary

all the computations necessary to achieve data compression of images. Two goals are the direct result of this initial demonstration by digital simulation. First is the goal of constructing an actual system based on the architecture of Fig. 5. This goal will be pursued (achievement of the goal being dependent upon the usual constraint, finding funds to build the device). A second goal is more encompassing. If optical computations for image data compression are feasible, the obvious action is to undertake the research necessary to raise the level of sophistication for optical data compression methods to that currently enjoyed by all-digital processes.

References

1. W. Pratt, Proc. Soc. Photo-Opt. Instrum. Eng. 74, 178 (1976).
2. A. Habibi and G. Robinson, IEEE Comput. Mag. 7, 22 (1974).
3. P. Wintz, Proc. IEEE 60, 809 (1972).
4. H. Whitehouse et al., Proc. Soc. Photo-Opt. Instrum. Eng. 119, (1977).
5. H. Federhen, presentation at SPIE conference, San Diego, August 1977.
6. A. Oppenheim and R. Schafer, Digital Signal Processing (Prentice-Hall, Englewood Cliffs, N.J., 1975).
7. J. O'Neal, Bell Syst. Tech. J. 45, 689 (1966).
8. N. Jayant, Proc. IEEE 62, 611 (1974).
9. S. Dashiell and A. Sawchuk, Proc. Soc. Photo-Opt. Eng. 74, 93 (1976).
10. S. Lee, Opt. Eng. 13, 196 (1974).
11. P. Tamura and J. Wyant, Proc. Soc. Photo-Opt. Instrum. Eng. 74, 57 (1976).
12. J. Goodman, Introduction to Fourier Optics (McGraw-Hill, New York, 1968).
13. A. Papoulis, An Introduction to Probability, Random Variables, and Stochastic Processes (McGraw-Hill, New York, 1965).
14. C. DeBoor, J. Approx. Theory 6, 50 (1972).
15. P. Nisenson, J. Feinleib, and S. Iwasa, Proc. Soc. Photo-Opt. Instrum. Eng. 45, 241 (1974).
16. J. Grinberg and E. Maron, Proc. Soc. Photo-Opt. Instrum. Eng. 118, 66 (1977).
17. H. Hou, "Least-Squares Image Restoration Using Spline Interpolation," USC report PIP-650 (U. Southern California, Los Angeles, 1976).

APPENDIX C

ANALYSIS OF FEEDBACK OPTICAL/VIDEO SYSTEMS

Analysis of Feedback
Optical/Video Systems

As mentioned in Section IV.2, fully parrallel DPCM would be an ideal component to process the correlation in image data, but such structures require coherent optical feedback and were discounted in previous research under this Grant. The basic problem of coherent optical feedback is phase ambiguity; optical wavelengths are very short compared to the dimensions of a feedback device, and the signal phase control must be equally precise <u>over the entire image plane</u>. We believe we have conceived a structure which will make it possible to combine the functional flexibility of optical feedback with the simplified phase control of electronic circuitry.

We wish to begin with a discussion of the general nature of two-dimensional feedback systems. Obviously, image feedback, of the sort used in a DPCM system, involves feedback of spatial and temporal information, but we will neglect temporal feedback to introduce the relevant concepts and then examine the associated temporal statility questions later.

A general structure for a two-dimensional feedback system is seen in Figure 1. H and G are two-dimensional Fourier transforms of the associated point-spread-functions, and R and C are the transforms of the corresponding input and output. It is obvious that the output of the processor can be written as:

$$C(u,v) = \frac{\gamma \alpha G(u,v)}{1 + \alpha \beta G(u,v)H(u,v)}R(u,v) \qquad (1)$$
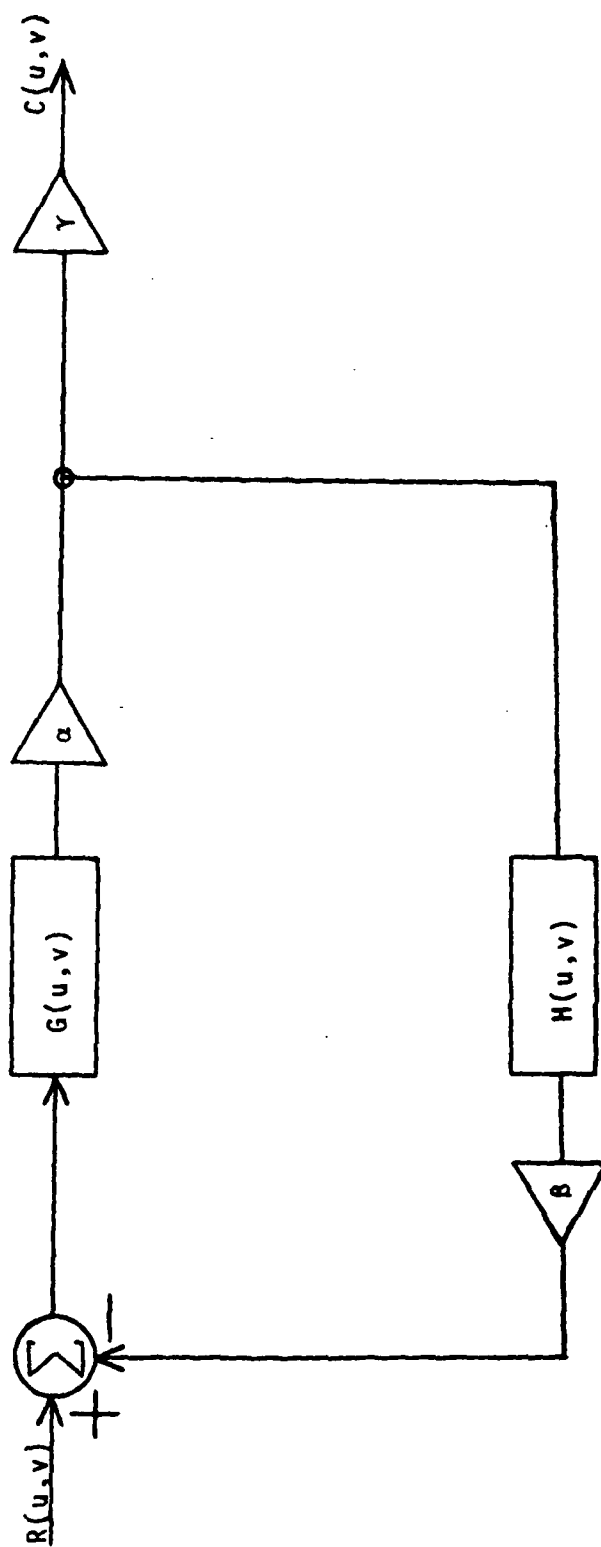
Figure 1

where α, β, and γ are constants that are either selectable and/or represent fixed gain constants of the loop elements. If α is set to unity and $g(x,y)$ is taken to be the dirac function $\delta(x,y)$, taking $\beta = \gamma$ the closed loop transfer function of the feedback processor becomes

$$\frac{1}{\frac{1}{\beta} + H(u,v)}$$

setting $H(u,v) = B(u,v)$ and letting β become arbitrarily large we obtain

$$C(u,v) = \frac{1}{B(u,v)} R(u,v) \quad . \tag{2}$$

On the other hand, for $\alpha = 1$, $\gamma = 1$, and $\beta = 0$, we have

$$C(u,v) = G(u,v)R(u,v) \quad . \tag{3}$$

Another response can be obtained by setting $\alpha = 1$, $\beta = \gamma$, $G(u,v) = B*(u,v)$ (where * indicates the complex conjugate), and $H(U,v) = B(u,v)$ whereby we obtain the transfer function

$$\frac{C(u,v)}{R(u,v)} = \frac{B*(u,v)}{\frac{1}{\beta} + |B(u,v)|^2} \tag{4}$$

which is recognized as the Wiener filter if β = SNR (the signal-to-noise ratio).

These two examples have been presented to illustrate the flexibility of a feedback synthesis. Recently Hausler and Lohman ([1] and [2]) have proposed the use of a closed loop TV system to create a feedback processor using incoherent light. In this research negative feedback was achieved by a combination of modulation techniques and an optical summation using a beam splitter. While sums are easily obtained optically, considerable care and expertise are required in implementing differences with incoherent

light, as Lohman's example illustrates. It would seem that the
subtraction step could be achieved electronically with greater
ease. Thus, one way of implementing the feedback processor of
Figure 1 would be as in Figure 2. Here r(x,y,t) is the system
input in the form of an electrical signal from a source such as a
TV camera. This signal is differenced with the feedback signal
through the video mixer. This difference signal is displayed as
a CRT image which is imaged by a vidicon or other camera with a
spatial filter corresponding to the point-spread-function g(x,y).
The spatial filter is realized by an apodization coating and con-
current defocus of the lens L_1 [3]. This obviously restricts the
transfer function G(u,v) to be one whose impulse is everywhere
positive. However, incoherent image "blur" functions have this
characteristic, so a large variety of realistic image processes
can be implemented. The feedback portion of the processor is ob-
tained through another CRT lens; with apodizing function and TV
camera combination as shown. The amplifiers with gains α, β, and
γ are used to represent the gains of various loop elements.

Some important characteristics of Figure 2 can be summarized
in the following points:

(1) All feedback effects operate with the video
(temporal , electronic) signal, hence the
problems of wavelength and phase references
are easily dealt with. For example, using
a conventional 30 frames per second, 525
lines-per-frame TV system and assuming tem-
poral bandwidth along a scan-line equivalent

Figure 2

40
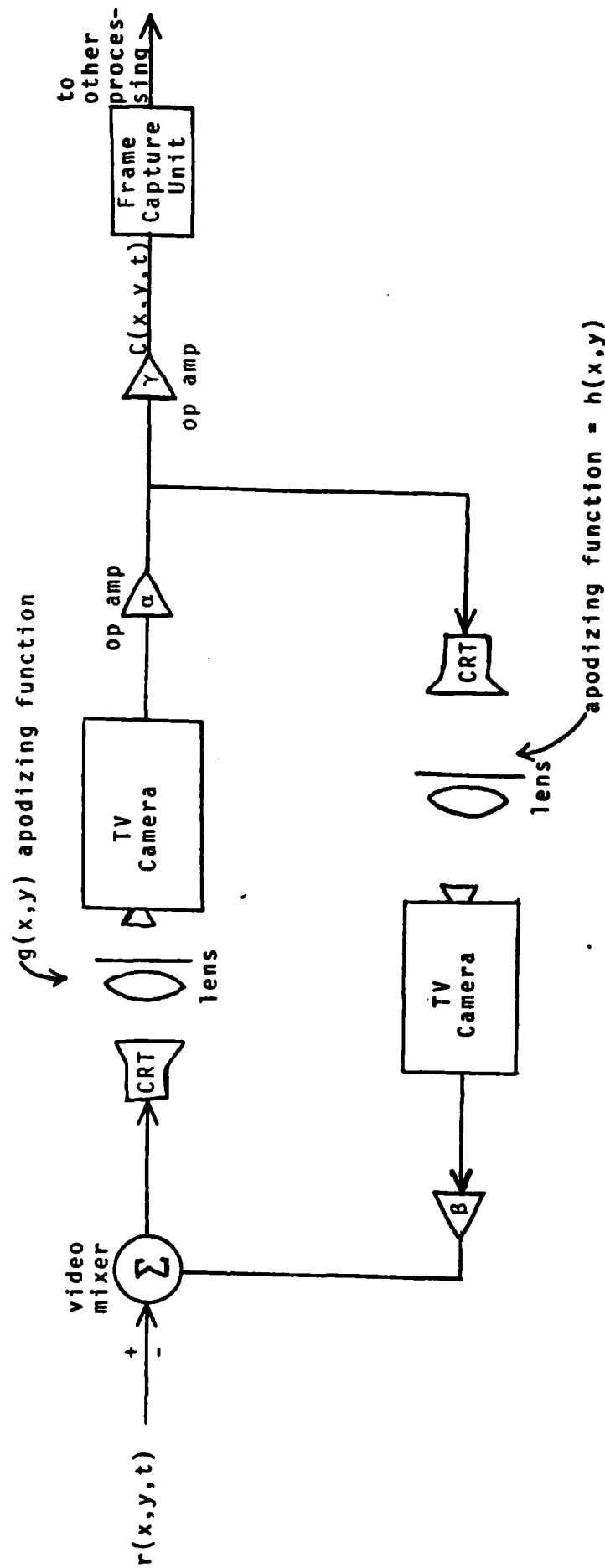
to 500 plus pixels, the nominal bandwidth of
the video mixing process would be 7.5 MHz.
Allowing sufficient additional bandwidth for
horizontal and vertical synch signals, blank-
ing pulse, D. C. restoration level, etc.,
gives a 10 MHz bandwidth, for which the wave-
length is 30 meters. Judicious choice of
laboratory set-up should keep the cabling in
the loop of Figure 2 to one meter or less;
hence, any wavelength and phase problems as-
sociated with the feedback shown in Figure 2
should be negligible.

(2) The temporal data in the video signals are
a means to encode x and y data. Since the
loop filters are spatial filters the result
of the combination of the temporal signals
and the spatial filters is an overall spatial
filter with an input-output response given
by Equation (1), where $G(u,v)$ and $H(u,v)$
are the two-dimensional frequency responses
of the point-spread-functions encoded by the
apodizing transparencies.

(3) The system shown in Figure 2 takes a video
signal as an input and yields a video signal
as output. Hence, the system could, at least
conceptually, be used to process video images
in a real time environment. Indeed Lohman

has described this sort of processor as an
iterative processor running at the TV frame
rate [2].

(4) The ultimate in flexibility and interaction
would be reached by using a programmable light
modulator for the apodizing transparency.
Combined with a servo on the focus control of
the vidicon, a great variety of optical trans-
fer functions could be rapidly inserted.

The concept of this feedback processor as an iterative algorithm
is intriguing in that it stimulates the question of the inter-
action between the temporal nature of the x and y data and raster
scanning of the x and y data. Furthermore, if one envisions the
processor of Figure 2 as a parallel processor in the x and y di-
rections then it effectively consists of a large number (depend-
ent upon the spatial resolution of the TV components) of temporal
feedback loops in which time stability is a question. In this
light we must devote some research effort into the temporal as
well as the spatial stability of the system of Figure 2. In
reality $r(x,y)$ is a function of time, $r(x,y,t)$, along with $c(x,y,t)$,
and $h(x,y,t)$. Time is a variable in $g(x,y)$ and $h(x,y)$ simply
because of the finite response times of the video signals. In
addition t is included as a variable in g and h to allow the flex-
ibility of tailoring the temporal response of the loop to insure
stable operation. Inclusion of the loop temporal characteristics
produces a relation between the output and input in the transform
domain as follows:

$$C(u,v,s) = \frac{\gamma\alpha\ G(u,v,s)}{1 + \alpha\beta\ G(u,v,s)H(u,v,s)}\ R(u,v,s) \tag{5}$$

where

$$G(u,v,s) = \iiint\limits_{-\infty\ 0}^{\infty\ \infty} g(x,y,t)e^{-j2\pi(ux+vy)}e^{-st}dxdydt \tag{6}$$

$$H(u,v,s) = \iiint\limits_{-\infty\ 0}^{\infty\ \infty} h(x,y,t)e^{-j2\pi(ux+vy)}e^{-st}dxdydt \tag{7}$$

and

$$R(u,v,s) = \iiint\limits_{-\infty\ 0}^{\infty\ \infty} r(x,y,t)e^{-j2\pi(ux+vy)}e^{-st}dxdydt \tag{8}$$

We recognize these expressions as two-dimensional Fourier transforms in the x and y directions with a Laplace transform along t. If we can make the assumption that $h(x,y,t)$ has a temporal response much faster than $f(x,y,t)$, $h(x,y,t)$ can be assumed to be independent of t so that $H(u,v,s) = H(u,v)$. Furthermore, if the TV camera in the forward loop can be chosen so as to have a temporal response (i.e., a finite rise time characteristic of some sort) independent of the spatial coordinates x and y, then $G(u,v,s) = G(u,v)\ G_1(s)$. Here $G_1(s)$ is the Laplace transform of the temporal characteristic of the forward path TV camera. The input-output relationship of Equation (1) becomes

$$C(u,v,s) = \frac{\gamma\alpha\ G(u,v)G_1(s)}{1 + \alpha\beta\ G_1(s)G(u,v)H(u,v)}\ R(u,v,s) \tag{9}$$

An interesting research question is the effect of a particular $G_1(s)$ upon the feedback processor's performance. Consider $r(x,y,t) = r(x,y)$ for $t > 0$ so that $R(u,v,s) = \frac{R(u,v)}{s}$. In addition if we

assume that $G_1(s) = \frac{k}{s+k}$ which would represent a simple low pass filter characteristic in the forward camera we obtain

$$C(u,v,s) = \frac{1}{s}\left\{\frac{\gamma\alpha\ G(u,v)}{\frac{s+k}{k} + \alpha\beta\ G(u,v)H(u,v)}\right\}R(u,v) \qquad (10)$$

Applivation of the final value theorem for Laplace transforms produces

$$\lim_{t\to\infty}(u,v,t) = \frac{\gamma\alpha\ G(u,v)}{1 + \alpha\beta\ G(u,v)H(u,v)}\ R(u,v) \qquad (11)$$

where

$$C(u,v,t) = \int\int\limits_{-\infty}^{\infty} C(x,y,t)e^{-j2\pi(ux+vy)}dxdy \qquad (12)$$

It is clearly seen that this is the same relationship as we had before in Equation (1). Because most cameras have a finite time lag between the image-induced charge being stored and the video signal output, some shift in the x direction is anticipated. This again raises the question of the separability of the raster scanning and finite time response of the TV camera. There is however, an extremely interesting way of circumventing this problem that deserves some attention. TV camera tubes such as the vidicon or the plumbicon can be used as image storage cubes. This is accomplished by focusing an image on the photo cathode while the scan is inhibited. The image induced charge will be stored (for a time dependent on the tube dark current) until the scan is initiated, where upon the image will be destructively read out as a video signal. By suitably blanking the CRT's and inhibiting the scan of the TV cameras in the system of Figure 2, we can achieve its sampled data counterpart as shown in Figure 3.

Figure 3

45
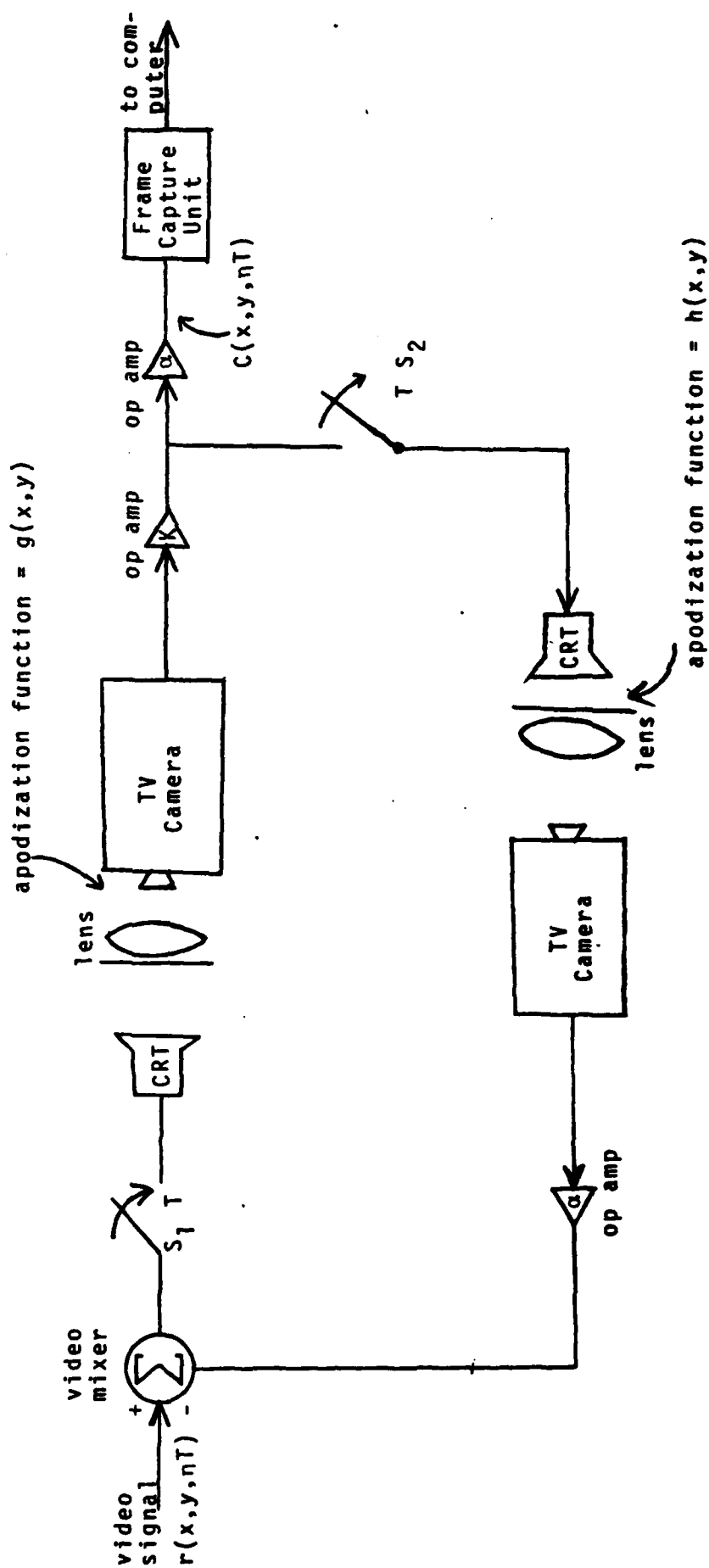
Operation of the feedback processor can be described as follows: closure of s_1 unblanks the forward path CRT and initiates the scan* of the feedback path TV camera. s_2 is open at this time which blanks the feedback CRT and inhibits the forward path TV camera scan. After completion of the feedback TV scan, the spatially filtered error signal is stored on the forward-path TV camera. s_1 is opened and s_2 is closed unblanking the feedback CRT (and initiating the forward TV camera scan). Upon completion of this scan, s_2 is opened and the spatially filtered output signal is stored on the feedback TV camera. This process is then repeated.

It is relatively straightforward to show that this scheme produces a sampled data feedback processor whose block diagram is given in Figure 4. The input-output relationship using a Z instead of a Laplace transform is as follows:

$$\underset{\sim}{C}(u,v,z) = \frac{\gamma\alpha \, Z^{-1}G(u,v)}{1 + \alpha\beta \, Z^{-1}G(u,v)H(u,v)} \, R(u,v,z) \qquad (13)$$

This produces the following recursion

$$\underset{\sim}{C}(u,v,nT) = G(u,v)\{\gamma\alpha\underset{\sim}{R}(u,v,(n-1)T) \\ - \alpha\beta H(u,v) \, \underset{\sim}{C}(u,v(n-1)T)\} \qquad (14)$$

where

$$\underset{\sim}{C}(u,v,nT) = \iint\limits_{-\infty}^{\infty} C(x,y,nT)e^{-j2\pi(ux+vy)}dxdy \qquad (15)$$

$$\underset{\sim}{R}(u,v,nT) = \iint\limits_{-\infty}^{\infty} r(x,y,nT)e^{-j2\pi(ux+vy)}dxdy \qquad (16)$$

---

*Scan means single frame scan since a single scan removes almost all of the stored charge in a plumbicon.
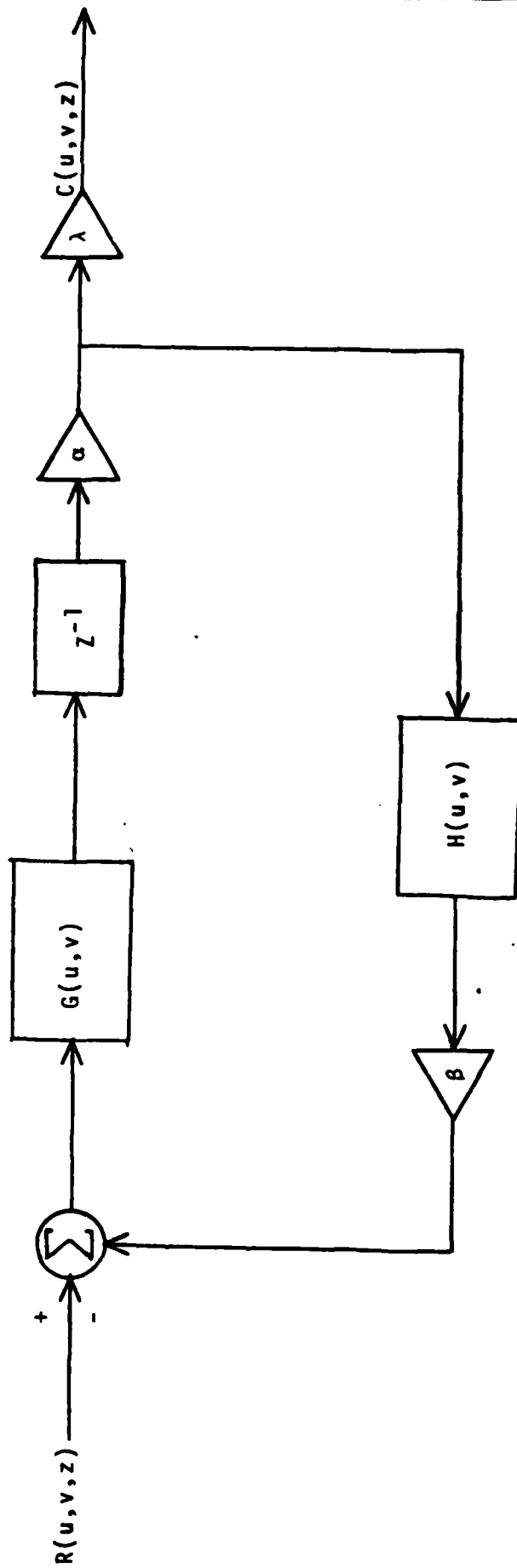
Figure 4

$$G(u,v) \Longleftrightarrow g(x,y)$$
$$H(u,v) \Longleftrightarrow h(x,y)$$

Again, assuming that $r(x,y,t) = r(x,y)$ for $t \geq 0$ we have $R(u,v,z)$ = $\frac{R(u,v)Z}{z-1}$, and application of the final value theorem for z transforms gives that

$$\underset{\sim}{C}(u,v,\infty) = \frac{\gamma\alpha\ G(u,v)}{1 + \alpha\beta\ B(u,v)H(u,v)}\ R(u,v) \qquad (17)$$

as before.

Thus we see that this formulation also produces the desired result. However, the optical readin and electrical readout functions in the TV cameras have been separated by the sampling functions simply by the fact that the image induced charge is allowed to accumulate prior to the initiation of the scan. This effectively eliminates the temporal characteristics of the x and y data due to raster scanning.

The utility of the incoherent optical structures we are discussing can be readily applied in a variety of ways. The applicability to image filtering is direct from the equations derived above. The applicability to DPCM is equally direct. A DPCM structure such as seen in Figure 5, would be readily implemented using the frame-storage techniques discussed in conjunction with Equations (13)-(17) above. Thus, N^2 parallel DPCM loops could be created by the optical/video hybrid components at a reasonable cost.

Figure 5

References

[1]   Hausler, G. and Lohmann, A., "Hybrid image processing with
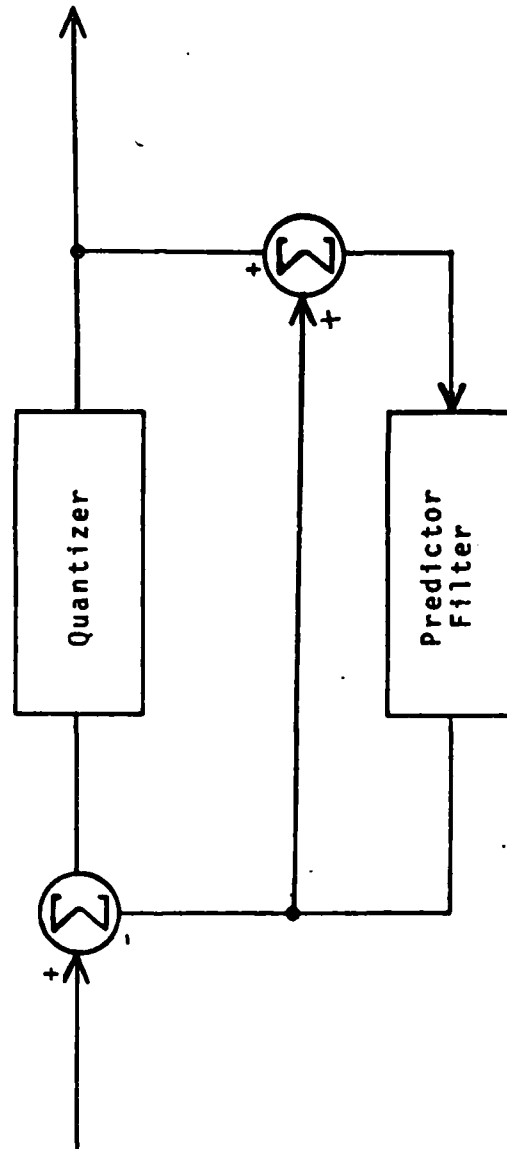   feedback", <u>Opt. Comm.</u>, Vol. 21, pp. 365-368, 1977.

[2]   Lohmann, A., "Suggestions for hybrid image processing", <u>Opt.
      Comm.</u>, Vol. 22, pp. 165-168, 1977.

[3]   Goodman, J., <u>Introduction to Fourier Optics</u>, McGraw-Hill,
      New York, 1968.

APPENDIX D

A TWO-CHANNEL MODEL OF IMAGE

PROCESSING IN THE HUMAN RETINA

A Two-Channel Model of Image Processing in the Human Retina

Douglas Granrath and B. R. Hunt

Digital Image Analysis Laboratory
Department of Systems and Industrial Engineering
University of Arizona
Tucson, Arizona 85721

Abstract

A model describing the decomposition of imagery in the human retina is developed based on the retina's cellular structure. Two types of retinal cells, horizontals and amacrines, perform spatial averaging across the retina to form a low-pass image channel. This low spatial frequency information is fed back to the retina's receptor cells to form a difference channel of high-passed spatial frequencies. Such a model is suggested by electro-physiological as will as psychophysical evidence. Analysis of the model characterizes the low-pass channel as a constrast channel and the difference channel as an edge detection channel. Application of the model to image quality assessment suggests a two factor approach involving metrics in the model's eye domain.

Introduction

Because of its great importance to our perception of the outside world, human vision has been seriously studied for centuries and is being studied today by many scientists in a variety of disciplines. Each line of vision research is aided by the dictates of its discipline, but in the process becomes limited by them as well. An experimental psychologist, for instance, studies vision by presenting light stimuli and recording the observers' responses, but he is then limited to input/output descriptions of visual behavior. The anatomist details the structures which comprise our visual system, but traditionally does not study their function. Function is studied by the physiologist, but resulting descriptions are limited by the enormous complexities involved. The need then exists for a discipline which can handle these complexities and provide theoretical descriptions to describe visual behavior on a more global level.

In its attempt to provide unifying descriptions of the complex machines man builds, engineering also provides theoretical tools to describe nature's complexities. Engineering models which aid in the design of image processing systems can be used to describe the human visual system as well. Analogies can then be made and components of our visual system identified which do image forming, sensing, coding, and transmitting, much like their physical counterparts. A good example of this is the vision model by Charles Hall and Ernest Hall[1] as shown in Figure 1. Each component in the model corresponds to a component or a process actually found in the eye: the low-pass filter models the ability of the eye's optics to form an image on the retina, the brightness function models the point-by-point transformation from light intensities to their neural representation, and the high-pass filter models the spatial interactions resulting from the cell connections in the retina. Such structural validity is a strength of the Hall and Hall model, but its retina model falls short in this regard. A spatially continuous model is valid for the low-pass filter since a continuous radiometric image is formed upon the retina, but the high-pass filter models the spatially discrete image processing which occurs in the retina's cells. Moreover, the retinal cells are fixed in their functional relationships with one another, and this structure imposes constraints on the image processing performed there.

```
  ──▶│ Low-Pass │──────▶│ Logarithmic │──────▶│ High-Pass │──▶
     │ Filter   │       │ Brightness  │       │ Filter    │
     └──────────┘       │ Function    │       └───────────┘
                        └─────────────┘
```

Figure 1. The Hall and Hall Model of Human Vision.

Cell Structure in the Human Retina

The retina is a network of nerve cells which are "hard wired" according to a particular organizing scheme. Although some of the connection parameters may differ in different areas of the retina and for other vertibrate species, the same basic scheme pervades the visual organization of the retina[2]. Figure 2 is a schematic representation of this organizing theme. The retina first senses the image with a layer of receptor cells. Two types of receptors, rods and cones, and three types of cones are found. Cones form our high

intensity or photopic system of color vision, while rods form our night vision or scotopic system which is achromatic. Just beyond the receptors is a thin layer of neurons called horizontal cells which connect receptor cells together from neighboring regions of the retina. A typical horizontal cell near the center of the retina (in the fovea) receives signals from seven cones; a foveal cone for its part feeds between two and four horizontal cells[3]. The horizontal cell then sends its signal out via a small number (approximately 4) of long, thin arms called axons to connect with an unknown number of receptor cells in neighboring regions. No horizontal-to-horizontal cell contacts have been documented in man[3], but they have been found in cat and rabbit retinas. The horizontal cell network is, therefore, probably not a continuous one in man, but considerable overlapping of its branches is in evidence.

Receptor Cells

Horizontal Cells

Bipolar Cells

Amacrine Cells
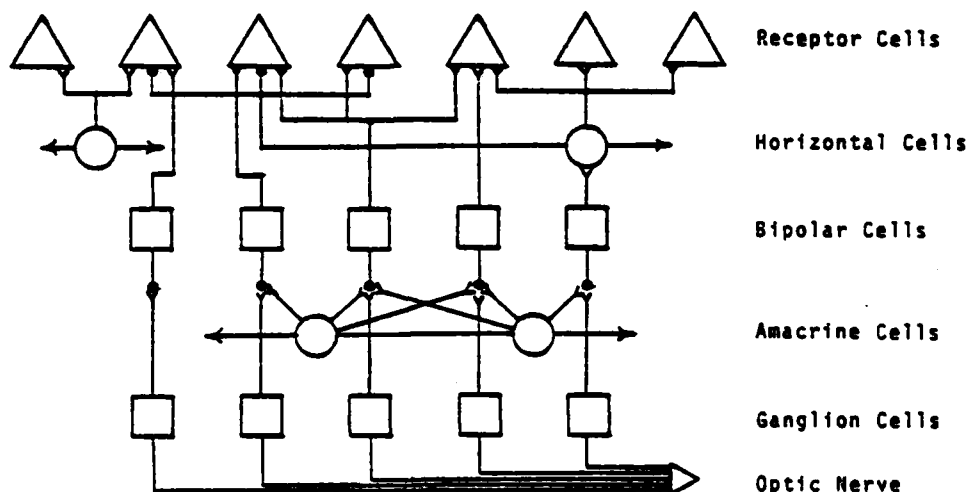
Ganglion Cells

Optic Nerve

Figure 2. Schematic Diagram of Retinal Cell Connections.

Forming the central layer of the retina are the bipolar cells. With single input and output arms of equal length, the bipolars connect the outer with the inner retinal layers, running parallel to the direction of light. Two types of bipolars connect with the cones in the fovea: midget and flat bipolar cells. A midget bipolar normally connects to one cone, while a flat bipolar connects to about seven cones. Each cone in the fovea connects to at least one midget bipolar and one or more flat bipolars. The midget bipolars thus appear to carry high-acuity image data and are likely color-coded according to their attached cones; the flat bipolars carry small-area averages which are probably monochromatic in nature[3].

The inner retinal layer is composed of amacrine and ganglion cells connecting to the output arms of the bipolar cells. Like horizontal cells, amacrines connect laterally across the retina, but unlike horizontals they form a functionally continuous network. The amacrines connect with bipolar axons in a way which allows for feedback to a bipolar from neighboring bipolars; amacrines also connect with one another and with ganglion cells. Approximately 10^6 ganglion cells form the final retinal layer; their very long axons bundle together to form the optic nerve, which contains all of the image information to be used by the brain.

The lateral connections of the horizontal and amacrine cells provide the basis for a two-channel decomposition of images. The horizontal cell network performs some degree of spatial averaging on the image at the first level of neural processing, thus forming a low spatial frequency version of the image. This information is fed back to bias the receptors, which make them respond relative to a local brightness average. The midget bipolar cells then transmit this high spatial frequency information with the local averages removed, while the flat bipolars transmit some form of the low spatial frequency information. In a similar manner the amacrine cell network provides for bipolar - bipolar feedback, and additional spatial processing may take place there. Amacrines appear quite complex in their function, providing for spatial-temporal interactions[4] and for image color encoding[5]. The two-channel organization of imagery is in evidence at the bipolar cell level, however, before the

intervention of the amacrines.

Two spatial channels are thus transmitted by the ganglion cells to higher areas in the brain. Two types of ganglion cells have, in fact, been observed: sustained and transient ganglion cells[6]. The sustained ganglions respond in a temporally steady manner to a constant stimulus and appear to be somewhat more numerous in the fovea than in the periphery of the retina. Transient ganglions, on the other hand, respond primarily to stimulus changes and exhibit decaying responses to steady state stimuli. As far as spatial frequencies go, the transient ganglions have been experimentally associated with low spatial frequencies and the sustained ganglions with middle to high spatial frequencies. Thus, the two channels which are formed by horizontal cell feedback to the receptors and are in evidence in the bipolar cells, are separately transmitted by the ganglion cells to higher visual centers in the brain.

Some recent findings point to a two-channel orgainzation of imagery in the brain's higher levels as well. In recordings from the foveal striate cortex of rhesus monkeys, Poggio, Doty, and Talbot[7] report two types of spatial frequency responses: band-pass and low-pass. They observe that low-frequency grating stimuli activate only particular low-pass neurons, depending on whether a contrast border falls within their receptive field. A band-pass neuron, on the other hand, requires a number of repeated edges (as in a grating) at an appropriate spatial frequency before it will respond, and at high frequencies only band-pass neurons are activated. At intermediate spatial frequencies of around 4 cycles/degree, significant numbers of both types of neurons are activated. All of the neurons they observed fall into one of these two categories, indicating a pervasive two-channel organization.

A Two-Channel Model of Spatial Interaction in the Human Retina

A two-channel model of retinal image encoding will now be expressed mathematically. The model will be two-dimensional, discrete in space, and achromatic. (Figure 3 gives a block diagram.) A discrete version of the continuous light intensity image which is optically formed on the retina constitutes the input image and is symbolized by the non-negative variable u_{ij}. The intensity-to-brightness mapping occurs first in the model and is performed in a point-by-point manner by:

$$x_{ij} = \log(1 + u_{ij}) \quad . \tag{1}$$

where x_{ij} represents the brightness image in the receptor cells. If we now define $y_{ij}^{(1)}$ and $y_{ij}^{(2)}$ to be the Channel 1 and Channel 2 output images in the optic nerve, the remainder of the model can be expressed in two discrete equations as follows:

$$y_{ij}^{(1)} = S\{K_1 \sum_k \sum_\ell h_{k,\ell}^{(1)} \cdot x_{i-k,j-\ell}\} \tag{2}$$

$$y_{ij}^{(2)} = \Gamma\{K_2 [\sum_m \sum_n h_{m,n}^{(2)} \cdot x_{i-m,j-n} - \sum_k \sum_\ell h_{k,\ell}^{(1)} \cdot x_{i-k,j-\ell}]\} \quad . \tag{3}$$

There are two unit sample responses or weighting functions indicated here: $h^{(1)}$ is a broad weighting function which models the relatively large inhibitory regions created by the horizontal cells, while $h^{(2)}$ is a narrow weighting function which models the smaller excitatory regions. Both weighting functions are Gaussian in form and are radially symmetric about the origin; more will be said about this choice later. The function $S\{\cdot\}$ is a spatial sampling of the low-frequency image in Channel1 of the model. This sampling is intended to model the likelihood that there are fewer connections for the low frequencies than for the high frequencies. The second channel of the model, Equation (3), is formed by the subtraction of the low-frequency image of Channel 1 from a high-frequency representation of the image. In modelling the retina at the fovea the high-frequency weighting function, $h^{(2)}$, would have essentially a one bipolar cell extent, and Channel 2 would reduce to:

$$y_{ij}^{(2)} = \Gamma\{K_2 [x_{ij} - \sum_k \sum_\ell h_{k,\ell}^{(1)} \cdot x_{i-k,j-\ell}]\} \quad . \tag{4}$$

A second nonlinearity is indicated in Channel 2 of the model by $\Gamma\{\cdot\}$; this is a saturation operator which models the bipolar cell's response limitations to receptor input. Finally, K_1 and K_2 are constants which model different channel gains.

Equations (1), (2), and (3) constitute a two-dimensional discrete model of neural image formation in the retina. The model has two nonlinearities, namely $\log(1 + u_{ij})$ and $\Gamma\{\cdot\}$. Because of this, it is important that the model be structurally valid, i.e., the nonlinearities must come at the correct places in the model. With linear, shift-invariant models structural validity is unimportant; the entire model could, in fact, be represented by a single transfer function without altering in any way the model's input/output behavior. Changing the location of a nonlinearity, however, changes the input/output behavior, so it

must be placed carefully for the proper overall model behavior. Since the two-channel model has been built with retinal structure in mind, it has been possible to incorporate the non-linearities in their proper locations.
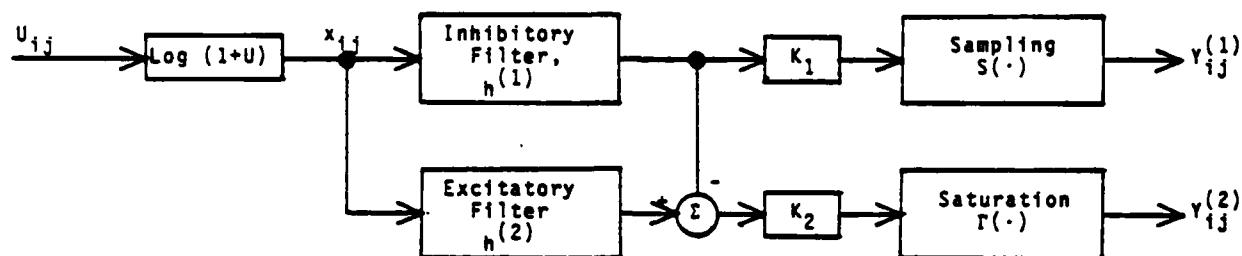

Figure 3. A Two-Channel Model of Spatial Interaction in the Human Retina.

The idea of using weighting functions to describe retinal behavior via spatial convolutions is not a new one. There has not been agreement, however, on the form the weighting function should take. A discussion comparing different types of weighting functions can be found in Macleod and Rosenfeld[8]. They prefer Gaussian type weighting functions and in particular a dual-Gaussian function of the form:

$$w(x) = \exp[-2x/s)^2] - (s/s') \exp[-(2x/s')^2] \tag{5}$$

The widths of the excitatory and the inhibitory regions can be adjusted separately here with the parameters s and s', but the positive and negative areas under the curve are kept equal to one another. Use of single Gaussian weighting functions for $h^{(1)}$ and $h^{(2)}$ in the two-channel model will result in a weighting function in the form of Equation (5) for the high-frequency channel of the model (Channel 2). In two dimensions the weighting functions will be radially symmetric (this is an approximation since the retina is somewhat anisotropic) and can then be expressed as:

$$h_{k,\ell}^{(1,2)} = \exp\{-(k^2 + \ell^2)/p_{1,2}^2\} \quad , \tag{6}$$

where the spread parameter, p_1 or p_2, uniquely specifies the function. Choice of the proper values for the model's parameters has been accomplished by computer simulation and comparison of the simulation results to experimental data. For foveal and near-foveal regions, comparable frequency response is obtained with $p_1 = 10$ and $p_2 = 0.5$, and a comparable reconstructed image is obtained with $K_1 = 1$, $K_2 = 3$, a Channel 1 sampling ratio of 16:1, and Channel 2 saturation levels of $\pm 1/6$ of the maximum image range in brightness. Details of this work can be found in the author's dissertation[9].

Choice of the logarithm for the intensity-to-brightness nonlinearity is in direct correspondence with Hall and Hall's model. Others have used the cube-root function[10], but this choice is not directly applicable here due to differences in model structure. Both choices for the brightness function yield sufficient dynamic range compression of intensity values, but the logarithm and the cube-root functions possess different non-linear characteristics which provide the basis for a choice. Two non-linear characteristics of the logarithm function will be explored in the following section on model analysis.

Analysis of the Two-Channel Model

The logarithmic brightness transform at the beginning of the two-channel model alters model behavior in a non-linear manner. In addition to global compression of the input data, the logarithm produces two effects of significance. The first of these is an automatic gain control of small amplitude signals, which enhances contrast at low light levels. The second effect occurs in conjunction with a multiplicative model of image formation and leads to a characterization of the model's two output channels. Mathematical treatment of both effects will be given in one continuous dimension.

Consider a small sinusoidal modulation of amplitude A at a relatively large constant intensity level, I, as the input:

$$u(x) = I + A\cos\omega x. \tag{7}$$

If contrast is defined by:

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \quad , \tag{8}$$

then the contrast of this input is simply:

$$C_{in} = \frac{A}{I} \quad . \tag{9}$$

After being log transformed, the input signal becomes:

$$f(x) = \log_b(1 + I + A\cos\omega x) ,$$ (10)

which can be expanded in a Taylor series about $(I + 1)$ as follows:

$$f(x) = \frac{1}{\ln b} \left[\ln(1 + I) + \frac{A \cdot \cos\omega w}{1 + I} - \frac{A^2\cos^2\omega w}{2(1+I)^2} + \ldots \right]$$ (11)

Since I is assumed large and A small, the first two terms constitute a good approximation to the output waveform:

$$f(x) = \frac{1}{\ln b} \left[\ln(1 + I) + \frac{A \cdot \cos\omega x}{1 + I} \right]$$ (12)

The contrast derived from this approximation, now, is:

$$C_{out} = \frac{\frac{A}{I + I}}{\ln(I + I)} = \frac{C_{in}}{\ln(I)}$$ (13)

a result which is independent of the original log transform's base, b. In comparison to the input contrast, the output contrast has been approximately reduced (with I large) by a factor of $(\ln I)^{-1}$. This means that the small-amplitude modulation of A_1 at an average intensity level of I_1 must be larger by a factor of $\ln(I_1 - I_0)$ in order to yield the same output contrast as a modulation of A_0 at a lower average intensity level, I_0, after both have been log transformed. In terms of input parameters, this condition can be expressed as:

$$C_1 = \ln(I_1 - I_0)C_0 ,$$ (14)

which relates input contrasts that are equivalent _after_ log transformation.

A logarithmic intensity/brightness transform provides global compression of the input image's dynamic range and boosts small signals at low average intensities relative to those at high average intensities. When acting in conjunction with the model's two-channel linear filter, the log transform has another interesting feature, namely the separation of object contrast from scene contrast. Analytical treatment of this effect requires a specific model of scene formation, but results of the analysis characterize Channel 1 of the model as a contrast channel and Channel 2 as an edge detection channel.

A basic assumption must first be made concerning the formation of a scene's radiance pattern. That is, the radiance from a scene is the product of two components, an illumination component and reflectance component. One-dimensionally this can be expressed by:

$$u(x) = u_r(x) \cdot u_i(x) ,$$ (15)

with the constraints:

$$0 \le u_i < \infty$$

and:

$$0 < u_{r,min} \le u_r \le u_{r,max} < 1.$$ (16)

Stockham[11] has argued for the validity of this model, and while setting $u_{r,min}$ equal to .005, points out that .01 is a likely minimum in virtually all situations. The important information in a scene is that which tells us something about the objects in the scene; this information is contained in the reflectance component of the radiation, u_r. Varying illumination across the scene combines with the object information in a multiplicative manner, making object detection difficult for a linear detector of radiation. An ideal object detector must somehow separate the reflectance from the illumination component in order to "see" the objects in a scene.

The two channels produced by the retina are able to separate the reflectance and illumination components if it can be additionally assumed that the illumination component is of lower spatial frequency content than the reflectance component. An analytical example will suffice to show this. Let,

$$u_r(x) = .5 + (.49)m_r\cos\omega_r x ; \quad 0 \le m_r \le 1,$$ (18)

be the reflectance of an object of spatial frequency ω_r and modulation m_r. The minimum reflectance possible here is .01 and the maximum is .99. Using the modulation definition of contrast given in Equation (8), the contrast of u_r is:

$$C_r = .98 m_r.$$ (19)

57

Now let,

$$u_i(x) = I + A\cos\omega_i x \ , \quad I > A \ . \tag{20}$$

be the scene illumination with a DC level of I and an amplitude of A; the contrast of u_i is:

$$C_i = \frac{A}{I} \ . \tag{21}$$

The following signal is thus taken as the input to the two-channel retina model:

$$u(x) = [.5 + (.49)m_r\cos\omega_r x] \cdot [I + A\cos\omega_i x] \ . \tag{22}$$

Let $f(x)$ be the log transformed signal:

$$f(x) = \log_{10}[1 + u(x)] , \tag{23}$$

which can be expanded in a Taylor series about $u(x)$:

$$f(x) = \log_{10}[u(x)] + \frac{1}{\ln 10}\left[\frac{1}{u(x)} - \frac{1}{2 \cdot u^2(x)} + \frac{1}{3 \cdot u^3(x)} - \dots\right] \tag{24}$$

Since the photopic region of vision is being modelled, $u(x)$ is large and the following approximation can be made:

$$f(x) \doteq \hat{f}(x) = \log_{10}[u(x)] \tag{25}$$

Substituting in for $u(x)$ gives,

$$\hat{f}(x) = \log_{10}[u_r(x) \cdot u_i(x)] \ , \tag{26}$$

or,

$$\hat{f}(x) = \log_{10}[u_r(x)] + \log_{10}[u_i(x)] \ . \tag{27}$$

With the input signal of Equation (22) this becomes:

$$\hat{f}(x) = \log_{10}[.5 + (.49)m_r\cos\omega_r x] + \log_{10}[I + A\cos\omega_i x] \ . \tag{28}$$

This signal next enters the linear portion of the model and produces a Channel 1 output, $y_1(x)$, and a Channel 2 output, $y_2(x)$. In order to find expressions for these two outputs in terms of their intervening transfer functions, the input, $\hat{f}(x)$, must be expressed as a sum of pure sinusoidal components. The second term of (28) can readily be approximated by the first two terms of its Taylor's expansion since I is much larger than A in the photopic region of vision. This gives:

$$\log_{10}[I + A\cos\omega_i x] \doteq \log_{10}(I) + \frac{A\cos\omega_i x}{I \cdot \ln(10)} \ . \tag{29}$$

More terms must be taken for an equivalent Taylor's approximation of the first term of (28):

$$\log_{10}[.5 + (.49)m_r\cos\omega_r x] \doteq \log_{10}[.5] + (.426)m_r\cos\omega_r x$$

$$- (.209)m_r^2\cos^2\omega_r x + (.068)m_r^3\cos^3\omega_r x - (.017)m_r^4\cos^4\omega_r x \tag{30}$$

Trigonometric substitutions are made to eliminate the powered cosine functions, yielding:

$$\log_{10}[.5 + (.49)m_r\cos\omega_r x] \doteq (-.301 - .104m_r^2 - .006m_r^4) + (.426m_r + .051m_r^3)\cos\omega_r x$$

$$- (104m_r^2 + .008m_r^4)\cos 2\omega_r x + .017m_r^3\cos 3\omega_r x - .002m_r^4\cos 4\omega_r x \ . \tag{31}$$

The five terms here include a DC term, a fundamental frequency term, and its first three harmonics, with each term a function of the reflectance modulation, m_r. With the addition of the two terms of Equation (29), a completed approximation of $\hat{f}(x)$ in terms of distinct spectral components is finally given by:

$$\hat{f}(x) \doteq [\log_{10}(I) - .301 - .104m_r^2 - .006m_r^4] + \frac{A}{I \cdot \ln(10)}\cos\omega_i x + [.426m_r + .051m_r^3]\cos 2\omega_r x$$

$$- [.104m_r^2 + .008m_r^4]\cos 2\omega_r x + .017m_r^3\cos 3\omega_r x - .002m_r^4\cos 4\omega_r x \ . \tag{32}$$

The log transform thus reduces the global intensity of the image (first term in (32)),

attenuates the illumination component (second term) by $[I \cdot \ln(10)]^{-1}$, and changes the reflectance component from a multiplicative to an additive component (third term), while also introducing harmonics (last three terms).

The approximated brightness signal, $\hat{f}(x)$, is next input to the eye model's linear filter which produces a low-pass output, $\hat{y}_1(x)$, and a high- or band-pass output, $\hat{y}_2(x)$. Let $G_1(\omega)$ and $G_2(\omega)$ be the modulation transfer functions of Channels 1 and 2, respectively. If it is assumed that $G_1(\omega)$ cuts off by ω_t with increasing ω and that $G_2(\omega)$ cuts off by ω_t with decreasing ω, then the outputs can be expressed by:

$$\hat{y}_1(\omega) = G_1(0)[\log_{10}(I) - .301 - .104m_r^2 - .006m_r^4] + G_1(\omega_1)[\frac{A}{I \cdot \ln(10)} \cos\omega_1 x] \ . \tag{33}$$

and,

$$\hat{y}_2(\omega) = G_2(\omega_r)[.426m_r + .051m_r^3]\cos\omega_r x - G_2(2\omega_r)[.104m_r^2 + .008m_r^4]\cos2\omega_r x$$

$$+ \ G_2(3\omega_r)[.017m_r^3]\cos3\omega_r x - G_2(4\omega_r)[.002m_r^4]\cos4\omega_r x \ . \tag{34}$$

An analytical conclusion can be made from Equations (33) and (34). Namely, the two-channel retina model acts to separate illumination intensity and contrast information into one channel and object reflectance information into the second channel. This is a desirable feature of a detector of objects to have, and it is apparently a design feature of our retinas.

Conclusions: Implications of the Two-Channel Model

The two-channel model describes how and suggests why the retina decomposes an image at the first levels of neural representation. The low-frequency channel allows the eye to represent a relatively large luminance range simultaneously (approximately three orders of magnitude). It also acts to bias the high-frequency channel, allowing the edge (high-acuity) data to be represented easily within the dynamic range of the bipolar and ganglion cells. The edge information gets a high priority in neural coding terms; more neural coding power goes to represent the high-frequency image compared to the low-frequency one. This implys that loss of image acuity and the edge information associated with it is viewed more critically by the retina than loss of global image contrast, which is represented mainly by the low-frequency channel. This has been the case in image quality assessment studies where acuity is consistently chosen as the primary quality factor with contrast the secondary factor[12].

An assessment of image quality can be made via the two-channel model. Image quality would be expressed with two factors: edge quality or degree of acuity and contrast quality. Experiments could be performed in which an original image is degraded by various combinations of blur and contrast to yield a set of test images. Quality measures could be computed for these test images by using a difference metric between their eye domain versions and an eye domain version of the original. Each test image could then have a pair of global quality measures for a given setting of the eye model (say, foveal) and a local pair for any desired subregion of the image. A factor analysis and a step-wise regression analysis could be performed between these metrics and results from a viewing experiment. The resultant would be a multiple regression equation predicting subjective image quality as a function of computable image metrics.

The two-channel model also implies that edge information is of fundamental importance to subsequent levels of neural processing in the brain. All of the image information must pass through the bipolar cell array, at which point the two image channels are already in evidence. The inner synaptic layer of the retina then receives these signals from the bipolars as input. This layer thus deals with edge images and contrast images, and any model at this level (a color coding model, for instance) must assume this. That edge information is part of the "language" of the retina has been shown experimentally. In a recent Scientific American article[13] experiments are described which show that if an edge separating an inner from an outer circle of two colors is eliminated by stabilizing the image on the retina then the color difference disappears too. Thus, the edges or borders in a scene are extracted from the image in the first layer of the retina and are basic to all of our visual perceptions of the world.

Acknowledgments

References

1. C. F. Hall and E. L. Hall, "A Nonlinear Model for the Spatial Characteristics of the Human Visual System", IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-7, No. 3, (March, 1977), 161-170.

2. S. L. Polyak, The Retina, Chicago: University of Chicago Press, 1941.

3. B. B. Boycott and J. E. Dowling, "Organization of the Primate Retina: Light Micro-scopy", Philosophical Transactions of the Royal Society B, Vol. 255, pp. 109-184.

4. F. S. Werblin, "Lateral Interactions at Inner Plexiform Layer of Vertebrate Retina: Antagonistic Response to Change", Science, 175, (1971), pp. 1008-1010.

5. H. Dawson, ed., The Eye, Volume 2A: Visual Function in Man, New York: Academic Press, 1976.

6. G. E. Legge, "Sustained and Transient Mechanisms in Human Vision: Temporal and Spatial Properties", Vision Research, 18, 69-81.

7. G. F. Poggio, R. W. Doty, Jr., and W. H. Talbot, "Foveal Striate Cortex of Behaving Monkey: Single-Neuron Responses to Square-Wave Gratings During Fixation of Gaze", J. of Neurophysiology, Vol. 40, No. 6, pp. 1369-1391, Nov., 1977.

8. I. Macleod and A. Rosenfeld, "The Visibility of Gratings: Spatial Frequency Channels or Bar-Detecting Units", Vision Research, 14, 909-915.

9. D. J. Granrath, "Models of Human Vision in Digital Image Bandwidth Compression", Ph.D Dissertaion, University of Arizona, August, 1979.

10. J. L. Mannos and D. L. Sakrison, "The Effects of a Visual Fidelity Criterion on the Encoding of Images", IEEE Transactions on Information Theory, Vol. IT-20, No. 4, (July, 1974), pp. 525-536.

11. T. G. Stockham, Jr., "Image Processing in the Context of a Visual Model", IEEE Pro-ceedings, 60, 7, (July, 1972), pp. 838-843.

12. H. Nyberg, "Multidimensional Scaling of Subjective Image Quality", FOA Report No. C 30039-H9, Nov., 1975, p. 28.

13. A. L. Gilchrist, "The Perception of Surface Blacks and Whites", Scientific American, Vol. 240, No. 3, (March, 1979), pp. 112-124.

APPENDIX E

AN IMAGE CODING ALGORITHM

USING SPLINE FUNCTIONS

Reprint from <u>SPIE Proceedings</u>, Vol. 149, 1978.

AN IMAGE CODING ALGORITHM USING SPLINE FUNCTIONS

Dennis G. McCaughey

Department of Systems & Industrial Engineering, University of Arizona
Tucson, Arizona 85721

Abstract

This paper presents an image coding algorithm using spline functions that is competitive with the more conventional orthogonal transform methods at data rates of 1 bit/pixel or less. Spline coding has the added attraction of an optical implementation arising from the fact that least squares image approximations also produces least squares approximations to the image derivatives. A first order spline is used to approximate the proper order derivative of the image whose order is determined by an analysis presented in the paper. The image derivative is then encoded and transmitted to the user who reconstructs the image by a k-1 order integration which can be done optically.

Introduction

This paper is concerned with the development of the concepts of image degrees of freedom and entropy from an approximation theoretic viewpoint for application to image coding. These concepts are used to develop a coding method using spline functions that can be implemented using optical processing techniques.

Treating the degrees of freedom of an image as approximation problem arises quite naturally in the context of image coding by transform methods, where an orthogonal transformation is performed on a sampled image matrix. A bandwidth reduction is obtained by transmitting only those transform coefficients above a certain threshold whose level is consistent with the desired error [1].

In this sense, the degrees of freedom of the image at an error of magnitude epsilon, or more succinctly, the epsilon degrees of freedom in terms of the orthogonal functions used, $DOF(\epsilon,\phi)$, is simply the number of functions in the set $\{\phi\}$ required to achieve the desired error, ϵ. The overall data rate $R(\phi,\epsilon)$ is the product of the number of bits, N_T, required to adequately represent the coefficients and the number of coefficients, $DOF(\epsilon,\phi)$ and is given by:

$$R(\phi,\epsilon) = N_T \cdot DOF(\epsilon,\phi)$$

Implicit in this is the assumption that the overall coding procedure can be separated into two parts: first, obtaining an adequate transform approximation of the image, and second, the quantization of the transform coefficients. By approaching the coding problem in this manner it becomes easier to understand one difficulty with the orthogonal transform coding methods. The large bandwidth reductions reported are due, in part to the compacting-of-image-energy property of orthogonal transformations. The difficulty in quantization of the coefficients is the result of the fact that any compacting in the transform domain is at the expense of an increased dynamic range in the transform coefficients, because of the conservation of energy inherent to all orthogonal transformations.

This would seem to indicate that a suitable set of transform functions would possess both good approximating properties, and transform coefficients whose dynamic range is of the order of that of the original image pixels. This idea could be extended to finding the best set $\{\phi\}$ minimizing $R(\phi,\epsilon)$ i.e.,

$$\min R(\phi,\epsilon) = \min(N_T DOF(\phi,\epsilon)) = R(\epsilon) = N_T DOF(\epsilon)$$

Again the assumption is made that the quantization and approximation steps can be separated.

Thus, one method of finding the epsilon-degrees-of-freedom, or the minimum data rate, would be to find the set of functions which, when used to approximate the image at an error rate epsilon, would require the fewest number of functions. This is a very difficult problem so the results in this work will be represented using k^{th} order splines. Splines are chosen due to their excellent image approximation properties [2], their desirable computation characteristics and the feasibility of an optical implementation.

Methods (Least Squares Spline Methods)

While the determination at each step of such a best approximating spline is simply a nonlinear minimization problem over the knots' defining the spline, it is computationally infeasible. Thus we must follow DeBoor [3] and settle for spline approximations with good, if not optimal, knot placements. In what follows, an easily implemented knot placement method will be given that can result in a significant error reduction over the uniform knot case. The results will be developed using splines, giving the following k^{th} order

spline approximation $\hat{f}(x,y)$:

$$f(x,y) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} S_{ij} N_{i,k}(\underline{\xi};x) N_{j,k}(\underline{\eta};y) = S_{k,N_x,N_y}(x,y) \tag{1}$$

where $N_{i,k}(\cdot;x)$ are the normalized B-splines of order k (degree k-1) described by DeBoor [4], and $\underline{\xi}$ and $\underline{\eta}$ are the knot vectors in the x and y directions respectively. The spline coefficients, S_{ij} are obtained by solving the following systems of equations:

$$f(x_\ell,y_m) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} S_{ij} N_{i,k}(\underline{\xi};x_\ell) N_{j,k}(\underline{\eta};y_m) \tag{2}$$

for $\ell=1,2,\dots N$ and $m=1,2,\dots N$. In matrix notation this becomes

$$[f(x_\ell,y_m)]_{N\times N} = [N_{i,k}(\underline{\xi};x_\ell)]_{N\times N_y}^{T} [S_{ij}]_{N_x,N_y} [N_{j,k}(\underline{\eta};y_m)]_{N_y \times N}$$

where $[\ ]^T$ indicates matrix transpose. To simplify notation let

$$[f(x_\ell y_m)] = [F]$$

$$[N]^{(\underline{\xi})} = [N_{i,k}(\underline{\xi};x)]$$

$$[N]^{(\underline{\eta})} = [N_{j,k}(\underline{\eta};y_m)].$$

Equation (2) becomes

$$[F] = [N]^{(\underline{\xi})^T} [S_{ij}] [N]^{(\underline{\eta})^T}$$

since $N>N_x$ and $N>N_y$, in general, equation (2) cannot be solved exactly. However, the spline coefficients that minimize the normalized least-squares error ϵ, given by the expression:

$$\epsilon = \frac{\sum\limits_{\ell=1}^{} \sum\limits_{m=1}^{} |f(x_\ell,y_m) - \hat{f}(x_\ell,y_m)|^2}{\sum\limits_{\ell=1}^{N} \sum\limits_{m=1}^{N} |f(x_\ell,y_m)|^2},$$

can be obtained by taking $[S_{ij}]$ to the

$$[S_{ij}] = [[N]^{(\underline{\xi})} [N]^{(\underline{\xi})^T}]^{-1} [N]^{(\underline{\xi})} [F] [N]^{(\eta)^T} [[N]^{(\eta)}_N (\eta)^T]^{-1} \tag{3}$$

The remainder of this subsection is concerned with the possibility of subsectioning the image and using different knot densities in each of the subsections, and with the quantization of the spline coefficients. It is reasonable that subsectioning might provide fruitful results, when one considers an L_2 error bound given by Schultz [5] for k^{th} order splines. Recalling that the error is given by the L_2 norm, $||\cdot||_2$ of the difference between the function and its approximation, this bound is given by

---

$\dagger$ The placement of the knots, the number of knots, and the number of spline coefficients are all equivalent Any one is sufficient to determine the others so that these forms will be used interchangeably in the sequel.

$$||f-\hat{f}||_2 = C(\bar{\rho})^{-k} \left|\left|\frac{\partial^k}{\partial x^k} f(x,y)\right|\right|_2 + \left|\left|\frac{\partial^k}{\partial x^{k/2}} \frac{\partial^{k/2}}{\partial y^{k/2}} f(x,y)\right|\right|_2$$

$$+ \left|\left|\frac{\partial^k}{\partial y^k} f(x,y)\right|\right|_2 \qquad (4)$$

where $\bar{\rho} = \max \{\max_i(\xi_{i+1}-\xi_i), \max(\eta_{j+1}-\eta_y)\}$

$\qquad C = O(4)$

$\qquad k =$ even integer

Thus if the image derivative energy is large only over a small region, then using a uniform knot k^{th} order spline with knot width equalling $\bar{\rho}$ as indicated by equation (4) should result in an overly good approximation of the image in those regions where the image derivative energy is low. Thus we should be able to obtain reasonable results by employing a different k^{th} order spline with uniformly spaced knots in each subsection if the knot density in each subsection is proportional to the value of

$$\left|\left|\frac{\partial^k}{\partial x^k} f(x,y)\right|\right|_2 + \left|\left|\frac{\partial^{k/2}}{\partial x^{k/2}} - \frac{\partial^{k/2}}{\partial y^{k/2}} f(x,y)\right|\right|_2 + \left|\left|\frac{\partial^k}{\partial y^k} f(x,y)\right|\right|_2$$

in that subsection.

After placement of knots and solution for the least-squares coefficients, the spline coefficients are uniformly quantized on a subsection by subsection basis. A uniform quantizer was chosen due to a lack of a better understanding of the coefficient statistics at this time. The number of coefficient quantization levels in each subsection was proportional to the variance of the spline coefficients in that subsection, with the maximum number of levels chosen for the subsection with the highest pixel variance. The proportionality constant which determines the number of quantization levels in each subsection is chosen to achieve the overall desired bit/pixel rate.

Since this is an adaptive quantization algorithm some overhead is necessary. The total number of bits required for transmission, N_T, is related to the total number of overhead bits, N_O, as follows and the number of coefficient bits N_R as follows:

$$N_T = N_R + N_O.$$

N_O can be determined by consideration of the fact that the overhead consists of the bits required to describe the subsection quantizers, N_q, the maximum possible number of coefficients per subsection, N_c, and the maximum possible number of bits per coefficient, N_b, in each subsection. Thus if N_s is the number of subsections the number of overhead bits is given by

$$N_O = N_s(N_q + N_c + N_b).$$

The number of bits required to describe the coefficients, N_R, is simply the sum over all the subsections of the number of bits required to describe the subsection spline coefficients.

Description of the subsection quantizers requires the maximum and minimum reconstruction levels. These are quantized to 32 bits each to ensure sufficient accuracy so that $N_q = 2 \times 32 = 64$. N_s is taken arbitrarily to be 64 so that a NxN image will have a maximum of 32 knots in the x and y directions if $N = 256$. Since the maximum possible number of coefficients $N_c = (32)^2$, 5 bits are sufficient to describe the number of knots, or equivalently, the number of coefficients. The maximum of quantization levels is taken to be 32 so that $N_b = 5$. These values are summarized in Table 1.

Experimental Results (Least Squares Spline)

To demonstrate the utility of using splines for image coding, an experiment was performed on the 256x256 pixel image shown in Figure 1. The image was partitioned into 64 subsections and approximated by second order splines. The unquantized spline approximation to the image is shown in Figure 2. The spline coefficients were then quantized at a rate of approximately 1 bit/pixel, including the overhead, and then used to produce the quantized image in Figure 3. The corresponding errors are shown in Table 2. Note that the quantization step at this bit rate has not introduced an excessive error increase over the unquantized spline

approximation. At 1.01 bit/pixel an error less than .5% is quite reasonable considering the non-optimal use of a uniform quantizer. A max quantizer [6] employing the proper statistical properties would most likely produce better results. Nevertheless, the visual qualities of the quantized reconstruction in Figure 3 are quite good and demonstrate that splines are a feasible approach to the image coding problem.

Optical Implementation (Derivative Spline)

The possibility of optically implementing the spline coding algorithms of the previous section is based on facts that: a least-squares k^{th} order spline approximation to an image produces a least-squares approximation to its derivatives up to order k-1, in terms of lower order splines and the divided differences of the spline coefficients; and that the k-1 derivative of a kth order spline is a first order spline of the form*:

$$S_{1,N_x,N_y}(x,y) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} C_{ij} \, N_{i,1}(\underline{\xi};x) \, N_{j,1}(\underline{\eta};y) \tag{5}$$

$$\text{where } N_{i,1}(\underline{\xi};x) = \begin{cases} 1 & \text{if } x \, \epsilon \, [\xi_i, \, \xi_{i+1}) \\ 0 & \text{otherwise} \end{cases}$$

$$N_{j,1}(\underline{\eta};y) = \begin{cases} 1 & \text{if } y \, \epsilon \, [\eta_j, \, \eta_{j+1}) \\ 0 & \text{otherwise} \end{cases}$$

An understanding of the process involved in obtaining a least-squares 1st order spline approximation of the k-1 derivative of f(x,y), $D_x^{k-1} D_y^{k-1} f(x,y)$, can be gained by consideration of figure (4). This figure shows the domain of definition of a particular subsection of $D_x^{k-1} D_y^{k-1} f(x,y)$, along with the knots defining $S_{1,N_x,N_y}(x,y)$. Since the least-squares approximation of $D_x^{k-1} D_y^{k-1} f(x,y)$ by a constant in the rectangle $[\xi_i, \xi_{i+1}) \times [\eta_j, \eta_{j+1})$ is obtained by setting the constant equal to the average value of $D_x^{k-1} D_y^{k-1} f(x,y)$ in that rectangle, C_{ij}^{j+1} is given by

$$C_{ij} = \frac{1}{(\eta_{j+1} - \eta_j)(\xi_{i+1} - \xi_i)} \int_{\xi_i}^{\xi_{i+1}} \int_{\eta_j}^{\eta_{j+1}} D_x^{k-1} D_y^{k-1} f(x,y) \, dx \, dy \tag{6}$$

with the estimate being

$$D_x^{k-1} D_y^{k-1} f(x,y) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \frac{1}{(\eta_{j+1} - \eta_j)(\xi_{i+1} - \xi_i)} \int_{\xi_i}^{\xi_{i+1}} \int_{\eta_j}^{\eta_{j+1}} D_x^{k-1} D_y^{k-1} f(\alpha,\beta) \, d\alpha \, d\beta \tag{7}$$

$$\times N_{i,1}(\underline{\xi},x) \, N_{j,1}(\underline{\ }, y)$$

$\hat{f}(x,y)$ is then obtained by a k-1 fold integration of equation (7) in the x and y directions with inclusion of the proper initial conditions. Since initial conditions in the higher order derivatives produce simple monomial changes in density across the face of the image (a situation which is unlikely to occur in practice) these are assumed to be zero. Thus only the coefficient matrix C and the image initial conditions need be quantized and transmitted to the user. The image is reconstructed by using an idealized coherent processor of figure (5). Here the approximation to $D_x^{k-1} D_y^{k-1} f(x,y)$ is the input. A filter whose transfer function corresponds to that of a k-1 order integrator in the x and y directions is placed in the back focal plane of Lens L_1. Thus the output in P_2 is the kth order spline approximation to f(x,y) minus the initial conditions. It should be noted that an actual implementation would involve the use of "leaky" integrators since the transfer function of an ideal integrator is unrealizable. However it is felt that this would not seriously affect the performance of the system. The initial condition estimates $\hat{f}(x,-1)$ and $\hat{f}(-1,y)$ are introduced with a beam splitter so that the final reconstructed image $S_{k,N_x,N_y}(x,y)$ appears in the output plane

* See the Appendix.

of Lens L_2.

Heretofore, the analysis has been idealized and simplistic in the sense that negative values of $S_{1,N_x,N_y}(x,y)$ exist and present problems to imaging devices that are intensity sensitive. This does not present an insurmountable difficulty in the reconstruction process since two processors can be implemented: one for the positive portions of $S_{1,N_x,N_y}(x,y)$ and one for the corresponding negative portions. Since $f(x,y)$ is always greater than zero the initial conditions need be included only in the processor for the positive portion. Labeling the appropriate portions,

$$S_{1,N_x,N_y}^{+}(x,y), \ S_{1,N_x,N_y}^{-}(x,y), \ S_{k,N_x,N_y}^{+}(x,y),$$

$$S_{k,N_x,N_y}^{-}(x,y), \ \text{and} \ S_{k,N_x,N_y}^{-}(x,y);$$

each can be obtained with a dual processor that includes a substraction step as shown schematically in figure (6). The subtraction step can be implemented either optically or electronically before a final image display step.

The difficulties with negative values of $S_{1,N_x,N_y}(x,y)$ are not so much with the reconstruction step, since a coherent processor can handle negative as well as positive values, but with the determination of $S_{1,N_x,N_y}(x,y)$ itself. $D^{k-1}D^{k-1}f(x,y)$ can be obtained optically but detecting its negative values with an intensity sensitive detector requires holographic recording techniques. This represents an unnecessary complication if this can be avoided by some other optical or hybrid processing technique. One such technique would involve imaging $f(x,y)$ with an NxN CCD camera whose output was the kth order divided difference of the pixel matrix F, $\nabla^{k-1}F[\nabla^{k-1}]^T$. Here

$$\nabla = \begin{bmatrix} -1 & 1 & 0 & - & - & - & 0 \\ 0 & -1 & 1 & 0 & - & - & 0 \\ & & & & & & \\ 0 & - & - & - & 0 & -1 & 1 & 0 \\ 0 & - & - & - & - & 0 & -1 & 1 \end{bmatrix}$$

and $F = [f(x_i,y_j)]$

$\nabla^{k-1}F[\nabla^{k-1}]^T$ is then averaged down to produce C by a microprocessor or a hard-wired algorithm. This is shown diagramatically in figure (7). The quantizer is shown incorporated in the averaging processor so that its output is the quantized version of C, C_q. The averaging rate is determined by the derivative energy processor. The rate information can be obtained either arithmetically from $\nabla^{k-1}F[\nabla^{k-1}]^T$ or optically from $f(x,y)$ - hence the two inputs shown. The output is the dimension of the C matrix or an equivalent quantity. C_q is shown as the output of the quantizer, where it is then split into two parts C_q^+ and C_q^-. where C_q^+ and C_q^- are given by

$$C_q^{+} = [\max (0, C_{ij}^{(q)})],$$

$$C_q^{-} = [|\min (0, C_{ij}^{(q)})|].$$

Thus C_q^+ and C_q^- are matrices consisting of non-negative elements such that

$$C_q = C_q^{+} - C_q^{-}.$$

Since C_q^+ and C_q^- are sufficient for the user to generate $S_{1,N_x,N_y}^{+}(x,y), S_{1,N_x,N_y}^{-}(x,y)$ and $S_{1,N_x,N_y}(x,y)$ these are transmitted to the user. Also transmitted are the quantized versions of the initial conditions $\hat{z}_q(x_i-1)+ f_q(-1,y_j)$.

Experimental Results (Derivative Spline)

To demonstrate the feasibility of the hybrid spline encoding system a simulation was performed, and these results were compared with those of the least squares spline at a data rate of approximately 2 bits/pixel.

The optically generated spline results are shown in figures 8 and 9. Figure 8 shows the unquantized results. A little blocking is evident, but otherwise the reconstruction possesses good detail, and is a generally faithful reproduction of the original. Figure 9 shows the results when the derivative spline coefficients are quantized at a rate of 2.04 bits/pixel including overhead. Here the detail in the mother and child has remained quite good with most of the features easily recognized. The background has been degraded and the blocking is more evident. The increased blocking is due to the suboptimal quantization of the initial conditions, to which the overall performance is quite sensitive. Nevertheless, the performance of this initial system compares favorably with the least squares spline results at 2 bits/pixel as shown in Table 3 and figure 10. As can be seen the error in the optical spline is elevated for both the unquantized and quantized versions. These come about due to the nature of the simulation. The results concerning the derivative properties of spline approximations as outlined in the appendix are true for the continuous model and not necessarily for the discrete case used in the simulation. Thus the actual optically implemented system might display a slightly improved performance. It is also expected that the use of the proper statistics in generating a Max quantizer would greatly improve the performance.

Conclusions

It would seem that spline functions are quite attractive for image coding purposes from both a performance and an implementation viewpoint. Concerning performance, an error of less than .5% at a rate of approximately 1 bit/pixel is certainly competitive with the orthogonal transform techniques. An optical implementation has been proposed and simulated that is both feasible to implement, and would provide a real time implementation. This optical implementation is a hybrid process since it combines a coherent optical processor for image construction, with a combination digital-noncoherent processor in the encoder.

It should be noted that the errors and corresponding data rate were achieved with a non-optimal quantizer and further work is necessary in this area. Further study and simulations are necessary and in progress to achieve a better understanding of the optical implementation. It is felt that these efforts should provide fruitful results.

Appendix

In this appendix, the properties of the derivatives of a kth order spline approximation to an image $f(x,y)$ are investigated. It will be shown that such an approximation also provides a least-squares approximation to the derivatives of $f(x,y)$ by taking the proper divided differences of the spline coefficients and expanding in terms of the proper lower order splines. In other words, if

$$S_{k,N_x,N_y}(x,y) = \underline{N}_k^T(\xi,x)\ S\ \underline{N}_k(\eta,y) \tag{A-1}$$

$$= \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} S_{ij} N_{i',k}(\xi,x)\ N_{j',k}(\eta,y)$$

is a least squares approximation to $f(x,y)$ on $[-1 \leq x, y \leq 1]$.
Then

$$D_x^i D_y^j\ S_{k,N_x N_y}(x,y) = \underline{N}_{k-1}^T(\xi,x)\ (\nabla^i S \nabla^{j^T})\underline{N}_{k-j}(\eta,x) \tag{A-2}$$

is least-squares approximation to $D_x^i D_y^j\ f(x,y)$ on $[-1 \leq x,y \leq 1]$.

Here

$$D_x D_y = \frac{\partial^i}{\partial x^i}\ \frac{\partial^j}{\partial y^j}$$

and ∇ is an N-1xN matrix, if $N_k(\xi,x)$ and $N_k(\eta,y)$ are NxN matrices ∇ is given by:

$$\nabla = \frac{1}{h} \begin{bmatrix} -1 & 1 & 0 & - & - & - & - & 0 \\ 0 & -1 & 1 & 0 & - & - & - & 0 \\ & \vdots & & & & & & \vdots \\ & \vdots & & & & & & \vdots \\ & \vdots & & & & & & \vdots \\ 0 & & & & & 0 & -1 & 1 \end{bmatrix}$$

where h is the knot mesh width of the spline.

The analysis will be performed in one dimension as the two dimensional equivalent result is obtained immediately using a direct product of splines.

The proposition is: If $S_{k,N}(x)$ is a least-squares approximation to $f(x)$ on the interval $[-1 \leq x \leq 1]$ given by

$$S_{k,N}(x) = \sum_{i=1}^{N} S_i N_{i,k}(\underline{\xi},x) \overset{\Delta}{=} \hat{f}(x)$$

$$= \underline{N}_k^T(\underline{\xi},x) \, \underline{S} \qquad\qquad (A-3)$$

then $S_{k-j,N}(x) = N_{k-j}(\underline{\xi},x) \nabla^j \underline{S}$ is a least-squares $k-j^{th}$ order spline approximation to $D_x^j f(x)$ on $[-1 \leq x \leq 1]$.

Proof:

If $S_{k,N}(x)$ is a least-squares approximation to $f(x)$ on $[-1 \leq x \leq 1]$ then the vector $\underline{S}$ must satisfy

$$\Gamma_k \underline{S} = \underline{f} \qquad\qquad (A-4)$$

where

$$\Gamma_k = \int_{-1}^{1} \underline{N}_k(\underline{\xi},x) \, \underline{N}_k^T(\underline{\xi},x) \, dx$$

$$\underline{f} = \int_{-1}^{1} f(x) \underline{N}_k(\underline{\xi},x) \, dx$$

so that

$$\underline{S} = \Gamma_k^{-1} \underline{f} \qquad\qquad (A-5)$$

Similarly, if $S_{k-j}(x)$ is a least-squares approximation to $D^j f(x)$, then $\nabla \underline{S} = \underline{Df}$ $\qquad (A-6)$

where

$$\underline{Df} = \int_{-1}^{1} \underline{N}_{k-1}(\underline{\xi},x) \, D_x f(x) \, dx \qquad\qquad (A-7)$$

so that

$$\nabla S = \Gamma_{k-1} \underline{Df}$$

It must be shown that $\underline{S}$ given by (A-5) implies (A-6). Therefore substituting the right side of (A-5) into the left side of (A-6) and writing $\underline{f}$ explicitly, one obtains

$$\Gamma_{k-1} \nabla \Gamma_k^{-1} \int_{-1}^{1} \underline{N}_{k-1}(\underline{\xi},z) \, \hat{f}(z) \, dz$$

$$= \int_{-1}^{1} N_{k-1}(\underline{\xi},x) \underline{N}_{k-1}(\underline{\xi},x) \nabla \Gamma_k^{-1} \int_{-1}^{1} \underline{N}_k(\underline{\xi},z) \, \hat{f}(z) \, dz \, dx$$

which is [4]:

$$\int_{-1}^{1} N_{k-1}(\underline{\xi},x) D_x \left[\underline{N}_k^T(\underline{\xi},x) \Gamma_k^{-1} \int_{-1}^{1} N_k(\underline{\xi},z) f(z) \, dz \right] dx.$$

But a least-squares estimate has the property $\hat{f}(z) = \underline{N}_k^T(\underline{\xi},z) \, \underline{a}$, for some $\underline{a}$ [7,8], so that

$$N_k^T(\underline{\xi},x) \Gamma_k^{-1} \int_{-1}^{1} \underline{N}_k(\underline{\xi},z) f(z) \, dz = \underline{N}_k^T(\underline{\xi},x) \underline{S} = \hat{f}(x)$$

giving a left side of (A-6) equal to

$$\int_{-1}^{1} \underline{N}_{k-1}(\xi, x) D_x \hat{f}(x)\, dx$$

$$= \underline{Df}^*$$

This result extends to higher order derivatives by induction, with a two-dimensional result obtained by a direct product of splines in the usual way.

References

1. Andrews, H.C., Computer Techniques in Image Processing, Academic Press, New York (1970).
2. McCaughey, D.G., and H.C. Andrews, "Variable Knot Splines for Image Approximation". IEEE Transactions on Computers (to appear).
3. DeBoor, C., "Good Approximations by Splines with Variable Knots," from Spline Functions and Approximation Theory, A. Meir and A. Sharma (eds.), Birkhauser Verlag Basel und Stuttgart (1973).
4. DeBoor, C., "On Calculating with B-Splines," Journal of Approximation Theory, Vol. 6, pp. 50-62 (1972).
5. Schultz, M.H., Spline Analysis, Prentice-Hall, Englewood Cliffs, N.J. (1973).
6. Max, J., "Quantizing for Minimum Distortion", IRE Transactions on Information Theory, Vol. ST6, pp. 7-12 (1960.
7. Luenberger, D.G., Optimization by Vector Space Methods, John Wiley and Sons, New York (1969).
8. McCaughey, D.G., and H.C. Andrews, The Continuous-Discrete Model Least Squares Inverses and Singular Function Expansions. Submitted to IEEE Transactions on Information Theory.

N_R	N_q	N_c	N_b	N_s	N	N_o	Rate bits/pixel
02012	64	5	5	64	256	4736	1.02

Table 1. Bit Allocation Summary

Figure (1). Original Image

Figure (2) Unquantized Least-Squares Spline Approximation

Figure (3) Quantized Least-Squares Spline Approximation

UNQUANTIZED SPLINE	QUANTIZED SPLINE
MSE = .230%	MSE = .381%

Table 2. Mean Squared Errors
for the Unquantized
and Quantized Splines
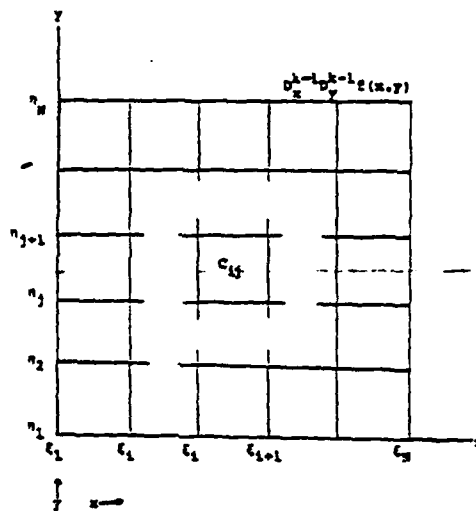at 1.02 bit/pixel
(including overhead)

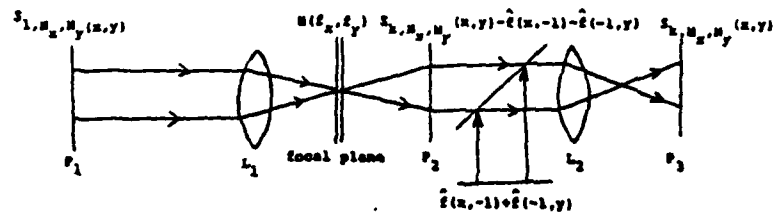Figure (4) Subsection geometry
for $S_{1,N_x,N_y}(x,y)$

Figure (5) Idealized coherent reconstruction
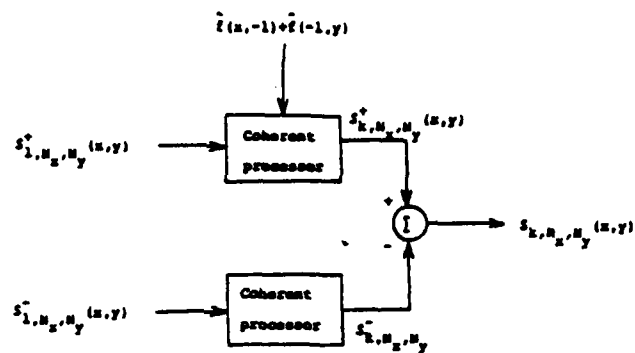of spline encoded image.

Figure (6)   Dual Processor of $S^+_{1,N_x,N_y}(x,y)$ and
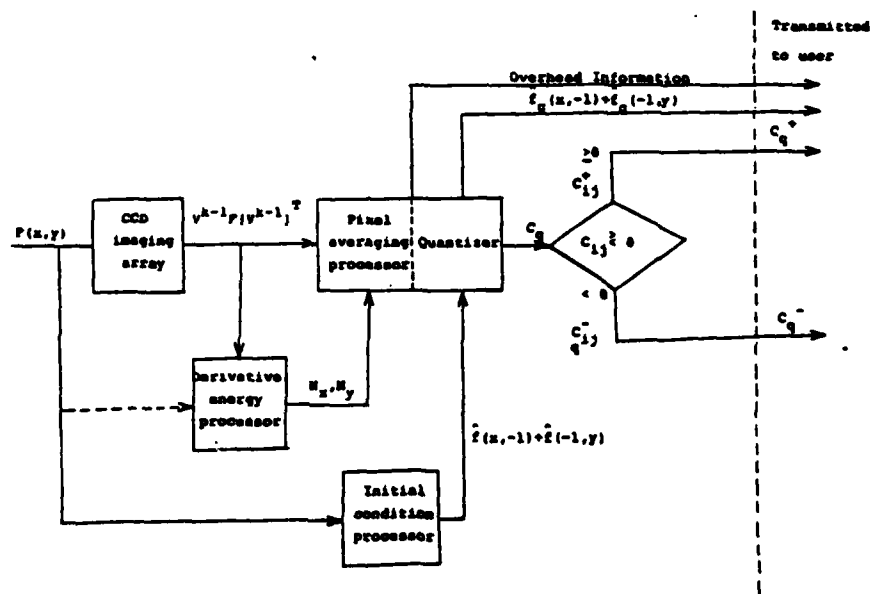$S^-_{1,N_x,N_y}(x,y)$ (both non-negative).

Figure (7).   Spline Encoding Processor

Figure (8). Unquantized   Figure (9). Quantized   Figure (10). Quantized
             Optical Spline   Optical Spline   Least-Squares
             Approximation   Approximation   Spline Approx-
                                                                           imation.

LEAST SQUARES SPLINE		DERIVATIVE SPLINE	
Unquantized		Unquantized	
	1.89 bits/pixel		2.04 bits/pixel
MSE = .230%	MSE = .234%	MSE = .468%	MSE = .639%

Table 3. Comparison of Results for
          Least Squares Spline with
          Those of the Derivative
          Spline at Approximately
          2 BITS/PIXEL.

Acknowledgements

APPENDIX F

OPTICAL IMPLEMENTATION OF A SPATIALLY

ADAPTIVE IMAGE DATA COMPRESSION SYSTEM*

Reprint from SPIE Proceedings, Vol. 232, 1979.

Optical Implementation of a Spatially Adaptive Image Data Compression System*

B. R. Hunt
Department of Systems Engineering
and Optical Sciences Center
University of Arizona
Tucson, Arizona 85721

Sergio D. Cabrera
Digital Image Analysis
Laboratory
University of Arizona
Tucson, Arizona 85721

Introduction

The increasing complexity and variety of image sensors has been the source of interest in the development of data compression for images. Image data has become one of the most active topics of research in digital image processing as a result [1]. The continued evolution of digital circuitry has caused the focus of data compression research to lie in digital implementations. However, there is also a potential for optical computations in image data compression, as was demonstrated in the concepts of interpolated DPCM [2]. The method of DPCM data compression is one of the most thoroughly studied techniques. DPCM achieves data compression by separating the image information into two parts: the low-spatial frequencies and the high-spatial frequencies. Low-spatial frequencies are re-tained by exploiting their predictability; high-spatial frequencies are retained at fewer significant bits, and substantial data compression is achieved. Interpolated DPCM is a mechanism for separating an image into low- and high- spatial frequency components, with a similar amount of data compression being achieved. The computations to achieve the sep-aration can be implemented by simple incoherent optical devices [2].

The greatest amount of data compression can be achieved only by adaptive processing, i.e., processing that changes as local image characteristics change [1]. Thus, the most efficient DPCM compression schemes are adaptive. The adaptive approach is not easily in-cluded in the optical processes examined for interpolated DPCM, however, since adaptation would require optical responses that vary within the image plane. To achieve adaptive data compression with optical components requires extremely sophisticated methods of adaptive optical spline interpolation [3]. The advantages of adaptive computation are thus offset by the extreme complexity required to implement them.

In the following we describe a system which has the capability of being partially adap-tive, but which is much less complex in system architecture and implementation. It can be considered as feasible for optical implementation with much less system complexity than the optical spline interpolation systems previously examined [3].

An Edge-Detecting Compression Technique

It has been known for some time that image data compression systems can be successfully developed around the separation of the original image into low- and high-frequency compo-nents. A fundamental consideration in such a system, therefore, is to answer the question: by what criterion do we define the separation between low- and high-spatial frequencies? The new data compression technique we describe below answers this question in terms of a criterion that we believe is relevant to the final quality of an image, that is a criterion concerning visible or perceived aspects of high-frequency information. We believe that the high-frequency information in an image is principally associated with the edges of objects in the original scene. Therefore, if we can successfully represent the edges of objects, we have the prospect of data compression that is acceptable in visual utility to the human being that constitutes the final end-user of the compressed data.

A schematic diagram of the proposed data compression system is seen in Figure 1. There are three separate parallel paths; structures to the left of the dotted line makeup the compression operations, and structures to the right of the line are the reconstruction oper-ations. The topmost of the parallel paths is responsible for deriving the low-spatial fre-quency information of the image which is to be compressed. This is done in a manner iden-tical to that developed for the interpolated DPCM technique previously discussed [2]. A set of subsamples are extracted from the original image. For example, if the original image possessed an intrinsic resolution equivalent to 512 x 512 pixels, the subsampled image might be 128 x 128, and would be created by extracting every 4th pixel of every 4th line. The subsamples are then quantized at a fixed number of bits and transmitted. At the re-ceiver the subsamples are used to create a low-frequency version of the original image. The subsamples are inserted into a matrix whose size is equal to the resolution in pixels of the original image, and zeroes are inserted in place of the pixels which were discarded in the subsampling process; e.g., for the case above every 4th pixel and line of the 128 x

128 subsamples would be inserted in a matrix of zeroes. This matrix of subsamples is then interpolated to replace the zeroes by interpolated subsample values.
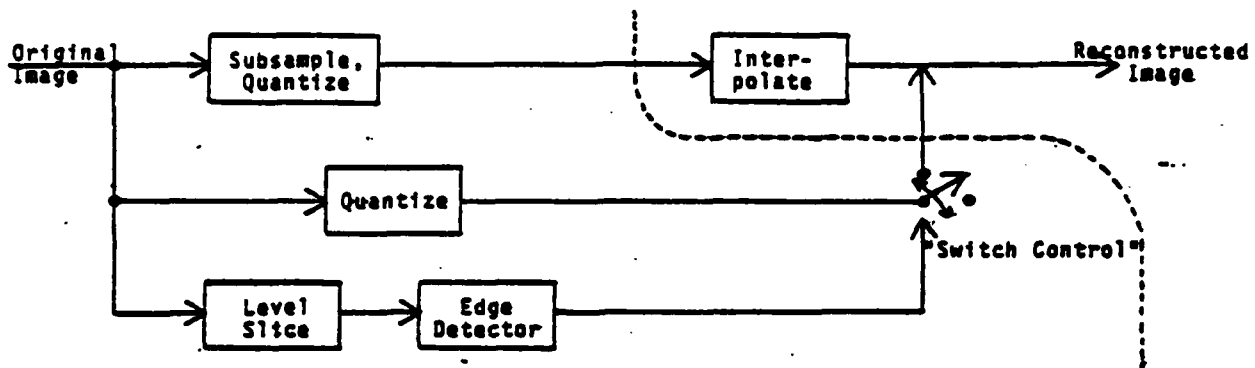

Figure 1: Compression System Schematic

As discussed in the paper describing interpolated DPCM, both the subsampling and interpolation operations can be carried-out by simple incoherent optical processes. For example, subsampling can be implemented with a simple focal-plane scanning mechanism, e.g., pick-off mirrors and apertures, and interpolation can be implemented via an out-of-focus apodized lens imaging a matrix of displayed subsamples.

The bottom two parallel paths in Figure 1 are the portions of the data compression system which represent the high-spatial frequency information in the original image. The bottom-most path consists of the mechanisms to detect edges. The rationale for the level slicing and edge detection is seen in Figure 2. A coarse slicing of levels causes a steep gradient to undergo a change in a number of levels in a short distance of space. An edge detector which can detect the existence of major level changes in a small spatial region has identified image areas which must be accurately represented to retain the sensitivity to edges of the human visual system. It is in this sense that the system of Figure 1 may be considered adaptive, i.e., it concentrates the high-frequency coding in regions of edges, edges being the most severe problems for high-frequency representation.
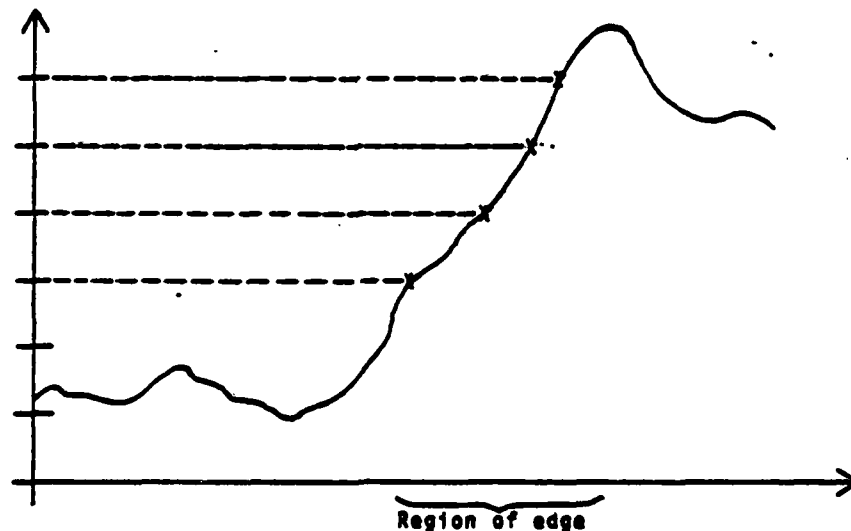

Region of edge

Figure 2: Adaptive Effects of Level Slice

The information concerning edge locations is used to "control" the switch in the middle path of Figure 1. The switch is not literal but symbolic of the following selection process. In regions where the low-frequency image dominates (that is, regions which are devoid of edges), it is assumed that the low-frequency representation created by the uppermost path is an adequate quality image. However, in regions where the edges are important (that is, regions where edges are detected) the actual value of image pixels are selected from the image region and transmitted. These literal edge pixel values are used to replace the

low-frequency representation of the edge pixels that would be constructed by the low-frequency interpolator. The system has the overall behavior of trying to represent the image by low-frequencies, but with adaptive mechanisms to select and retain the most important high-frequency content for the edge structures.

The implementation of the bottommost path in Figure 2 can also be achieved by optical processing, or in combination with discrete focal plane sensors. For example, the level slicing operation has been demonstrated by the nonlinear halftone work of Sawchuk [4]. The edge detection operation can be broken in two steps: edge calculations and edge thresholds. The edge calculation is, in essence, the convolution of the level-sliced image with a particular point-spread-function, and a convolution can be implemented optically [5]. The edge threshold is a trivial case of level-slicing and can be implemented, again, by the Sawchuk method. Alternatively, the threshold can be calculated digitally. In this case, the image output from the level slice would be sampled with a discrete sensor array, e.g., an array of silicon photo-diodes or imaging CCD elements. A simple digital processor would access the sensor array elements, perform analog-to-digital conversions and digitally implement the threshold operations.

The optical implementation of the level-slice and edge calculation would require a step in converting the incoming image from non-coherent to coherent illumination, since the proposed optical processors require coherent sources. There are a number of ways of achieving this conversion, however, and we will not dwell upon the specifics.

The major computations for data compression in Figure 1 can be implemented optically. The actual transmission of compressed image values would most likely be in digital format, given the increasing preference for digital data transmission techniques. Thus, following the computations for data compression the compressed data must be coded for transmission, and digital hardware will be employed. The final system is, thus, most appropriately a hybrid optical/digital system. In the following section we indicate some of the digital coding considerations.

Simulation of the Adaptive System

To assess the efficiacy of an edge-detecting optical compression system, a simulation of the architecture in Figure 1 was carried out with digital image processing techniques, that is, optical convolutions for interpolation were implemented by digital convolutions, optical level-slicing was implemented with a digital requantization, etc.

In the simulation of the edge-detection compression on an image, the following parameters are important:

(1) Subsampling increment, the spacing between pixels retained for reconstruction of a low-frequency image;

(2) Number of slice levels, the number of uniform steps into which the pixel radiance values are divided;

(3) Edge threshold measure, the number of steps in level which are chosen to represent an edge;

(4) Subsampling quantization, the number of bits to quantize a sub-sample for transmission;

(5) Edge quantization, the number of bits to quantize a pixel chosen as an edge for transmission;

(6) Edge coding, the method chosen to represent the location of detected edges.

Ideally, the optimum values for these parameters would be chosen on the basis of a comprehensive theory of the system. No such theory has been developed, however, and we have initially chosen values for the parameters from experimentation with the method.

Figure 3 shows an original digital image, sampled at 256 x 256 pixels, with each pixel quantized at 8 bits. For the simulation of the compression system, the following parameters were chosen:

(1) The subsampling increment was every 4th line and every 4th pixel. The image interpolated at the receiver from this 4 x 4 subsampling is shown in Figure 4. The loss of resolution and the inability to identify objects in this image is pronounced. The sub-sampling quantization was 4 bits, i.e., the interpolation

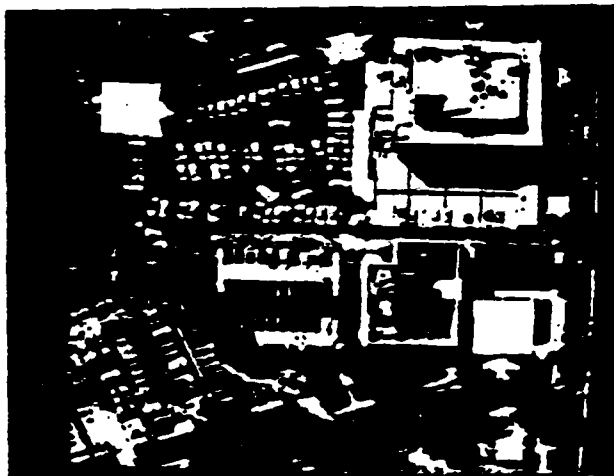shown in Figure 4 was created from 4-bit requantization of the subsamples selected from Figure 3.
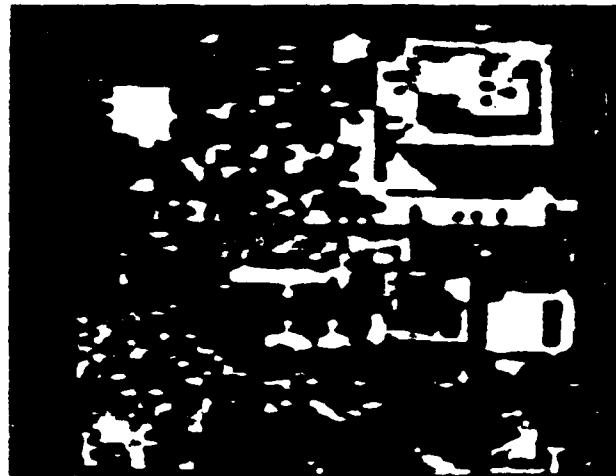

Figure 3:   Original Image (256 x 256)

Figure 4:   Reinterpolation from 4 x 4
Subsampling of Figure 3.

(2) For level-slicing, 16 slices were chosen, the 16 levels being uniformly
distributed between the minimum and maximum pixel grey values (0 and 255).

(3) An edge was assumed to be 5 or more steps in the slices produced by the
level slicing. Figure 5 shows the map of edges produced by using an edge-
detector with this 5-step threshold criterion.

(4) Each pixel selected as an edge was quantized at 4 bits. The result of
inserting these 4-bit edge pixels into Figure 4, at the edge locations
defined by Figure 5, is shown in Figure 6, the reconstructed image.
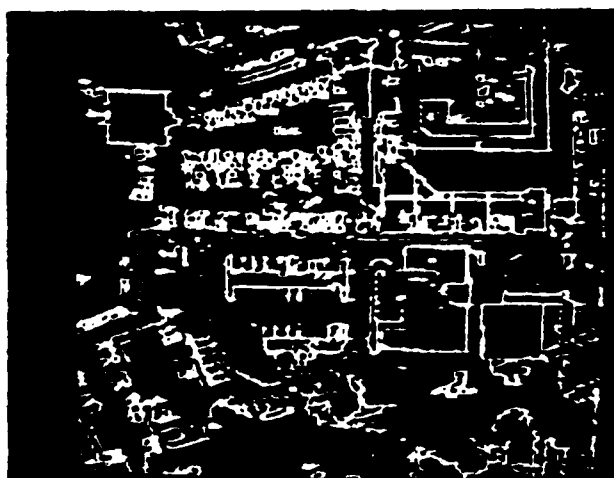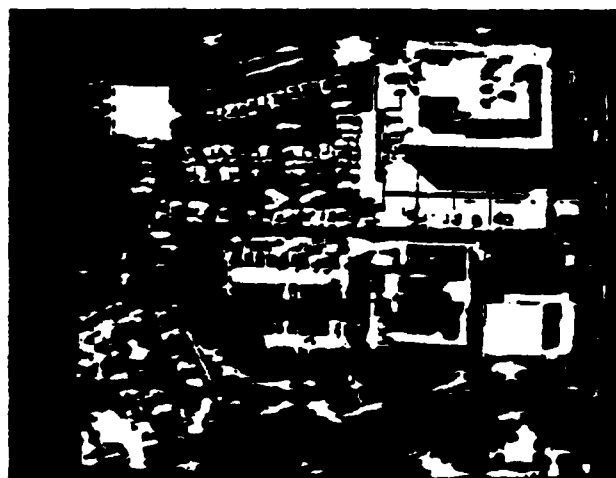

Figure 5:   Edge Map for Figure 3

Figure 6:   Reconstruction: Bit Rate =
1.385 bits/pixel

Two processes critical to the system operation are the edge detector and the coding of
edge locations. A number of edge detectors have been experimentally evaluated, including
some of the most commonly found in current literature, e.g., Laplacian operator, Sobel

operator, [6] etc. Experience with these different edge operators, and others devised for the simulation, has shown that there is little overall difference in compression performance among a variety of choices for the edge operator. Consequently, the operator which is most direct or simple to implement is the rationale for choosing between the different edge operators. The coding of edge locations does not offer as many options, however. It is necessary to represent the location of an edge in as few bits as possible, in order that the bits saved by the edge-detecting process not be lost in transmitting edge locations. This also requires a compromise in ease of implementation. A simple, and yet relatively efficient technique is run-length coding [6,7], coding the length of a run, where a run is defined to be the number of adjacent pixels which are alike, that is, either edge or non-edge pixels. Since edges are less frequent in occurence than non-edges, a shorter code word is required for edges than non-edges. In Figure 6 a run-length code of 2-bits for edges and 5-bits for non-edges was chosen.

The compression efficiency achieved in the creation of Figure 6 is summarized as follows. With the 4 x 4 subsampling at 4 bits per sample, a total of .25 bits per pixel are required. The run-length code for edge locations in Figure 6 amounted to 5211 edge runs and 5371 non-edge runs, and the total code bits in the run-length code is equal to 0.5688 bits per pixel. Finally, there are 9276 detected edge values in Figure 6, and quantizing these edge values at 4 bits each leads to a total of 0.5661 bits per pixel. The total bit requirements to represent Figure 6 is, thus, 1.385 bits per pixel. Note that in each of the above items, "bits per pixel" refers to bits per pixel *of the original image* (not per subsampled or edge pixel). Thus, Figure 6 represents a bit reduction from 8.0 bits per pixel to 1.385 bits per pixel.

Figure 7 and Figure 8 represent another pair of images, sampled at 512 x 512 pixels and processed with parameter choices identical to Figure 3. The resulting bit rate in Figure 8 is 0.922 bits per pixel, as compared to 8 bits per pixel in the original of Figure 7.
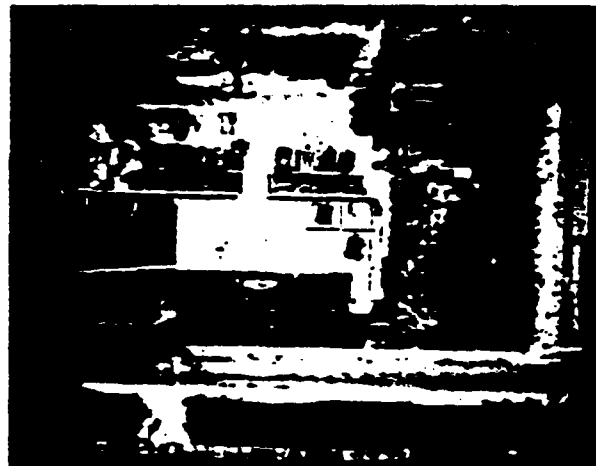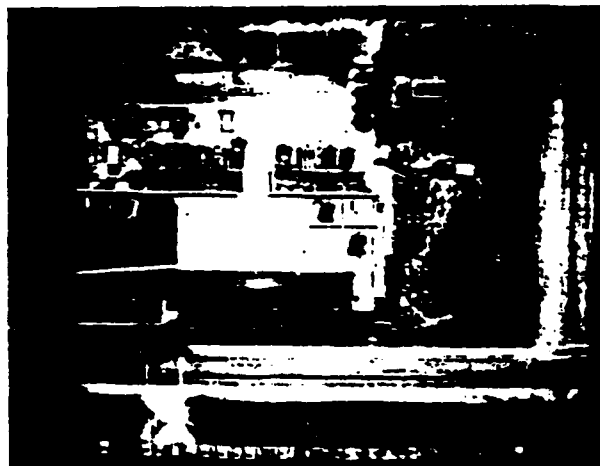

Figure 7: Original Image (512 x 512)   Figure 8: Reconstruction: Bit Rate = 0.922 bits/pixel

The visual properties of both Figures 6 and 8 is similar, in that low-contrast non-structured details present in the original image are replaced, through the subsampling/interpolation process, by blurred shapes. The edges that define and outline objects are retained, however, and all major objects visible in the original are visible in the compressed and reconstructed image. There are some visible artifacts at the edges of the reconstructed images.

Summary

We have described an image data compression system which uses edge-detecting mechanisms to adaptively code image features to which the eye is most sensitive. There are some obvious directions in which future research in this method should be pointed. For example, the use of a physiological model of edge-detection would assure that edge detection is sensitive to properties to which the human eye is also sensitive. A general model of compression performance, as a function of the system parameters discussed above, is also

needed, in order to optimize the overall system performance for minimum bit-rate and maximum image quality. A general goal for this type of system is to achieve an average bit-rate of 1.0 bits per pixel with image quality virtually identical to the original image. Whereas this goal was not achievable with the simpler method of interpolated DPCM, we believe the method discussed herein may offer the potential to achieve this goal in optical componentry for image data compression.

References

1. Hunt, B. R., "Digital Image Processing", *IEEE Proceedings*, Vol. 63, pp. 693-708, 1975.
2. Hunt, B. R., "Optical Computing for Image Bandwidth Compression: Analysis and Simulation", *Applied Optics*, Vol. 17, pp. 2944-2951, 1978.
3. McCaughey, D. G., "An Image Coding Algorithm Using Spline Functions", *SPIE Proceedings*, Vol. 149, 1978.
4. Dashiell, S. and A. Sawchuk, "Image Processing by Nonlinear Halftone Screens", *SPIE Proceedings*, Vol. 74, pp. 93-98, 1976.
5. Goodman, J. W., *An Introduction to Fourier Optics*, McGraw-Hill, New York, 1976.
6. Pratt, W. K., *Digital Image Processing*, Wiley, New York, 1978.
7. Gonzalez, R. C. and P. Wintz, *Digital Image Processing*, Addison-Wesley, Reading, Massachusetts, 1977.

APPENDIX G

A HYBRID OPTICAL/DIGITAL INTERFRAME

IMAGE DATA COMPRESSION SCHEME

A Hybrid Optical/Digital Interframe
Image Data Compression. Scheme

B. R. Hunt and H. Ito

Department of Systems Engineering
University of Arizona
Tucson, Arizona 85721

Abstract

Image data compression methods have been dominated by digital computations. In this paper we discuss a data compression concept which employs optical computations as part of the compression process. Simple optical processes are used to separate an image into low frequency and high frequency components. These components are then subjected to temporal compression, for multiframe imagery, by using a DPCM frame-buffer structure. Simulations of the process are shown, with reasonable performance being seen at multiple frame compression rates of 1.75 bits per pixel.

Introduction

Image data compression is a very active topic of research in image processing. This is not surprising when one considers the extensive sources of imagery currently being used for a variety of purposes, e.g., LANDSAT imagery, medical imaging, non-destructive testing, etc. Many such imagery situations require either the point-to-point transmission of imagery and/ or the archival storage of imagery. Given the economic costs associated with transmission or archival storage of masses of image data, the desirability of image data compression to reduce these costs is obvious.

Image data compression schemes have been dominated by digital processes. That is, the required computations for an image data compression scheme have been implemented by digital processes. This is not surprising, given the emphasis on compression schemes which require the inherent flexibility of a digital scheme, e.g., adaptive cosine transform compression. However, an investigation of different architectures for image data compression can reveal feasible data compression methods for which optical computations can replace digital computations. The successful discovery of such architectures is interesting because they would represent an extension of the repertoire of optical processing functions into new situations where optical and digital processes would be directly competitive.

One such success in the search for image data compression architectures with optical implementation is the IDPCM method [1]. This is a data compression scheme which functions analogously to conventional digital DPCM compression [2], except that the specific compression steps are implemented by optical processes.

Interframe data compression methods are applicable only to imagery sources which are temporal in variation, e.g., the successive frames of a commercial broadcast television signal. Obviously, if the extent of changes in successive frames of a multiframe sequence is small, then there will be great temporal redundancy of the spatial information recorded in the image frames. This is the purpose of an interframe compression system: to remove the temporal redundancy of the imagery's spatial data.

In the following article we examine a data compression system which combines optical spatial processing with temporal processing for interframe data compression.

Interframe Architecture

The basic structure which we propose for the hybrid digital/optical interframe compression system is shown in Figure 1. The upper portion of the figure represents the portion of the system responsible for data compression, whereas the lower portion is the reconstruction system. The portions of the system using digital and optical componentry are clearly segregated, as well. We will discuss the overall operation of the schematic and then describe the actual componentry implementation of the various blocks in the diagram.

To the left of the nodes marked ① in Figure 1 is the portion of the system where spatial data redundancy is eliminated. The image is subsampled, and the subsamples are used to reconstruct a low-frequency version of the original image using bilinear interpolation of the subsamples. The low-frequency version of the image is then subtracted from the original image. This is, of course, equivalent to a high-pass filtering of the original image.

To the right of the nodes marked ① in Figure 1 is the portion of tne system where temporal data reoundancy is eliminated. The quantizer and feedback structure are identifiable as similar to a conventional DPCM image data compression system [3]. However, tnere is one great difference between the structure in Figure 1 and conventional DPCM data compression: parallel vs. serial data flow. In a conventional DPCM system a set of successive sequential samples are extracted from the image and saved in the data buffer for prediction and differencing with succeeding image samples. However, in Figure 1 it is successive frames of imagery which are buffered, used for prediction, and differenced. Thus, the architecture of Figure 1 signifies parallel flow of image pixels around the quanitzation/feedback loop, as well as parallel differencing at the nodes marked ① . Rather than specific samples from an image, the DPCM loop in Figure 1 represents parallel image-plane to image-plane operations. It is equivalent to a bank of N x N serial DPCM processors operating in parallel (where N x N is the pixel resolution of the image plane).

Likewise, the nodes marked ② in Figure 1 represent similar parallel operations in the reconstruction process. That is, a frame buffer saves each successive frame and sums it, in parallel with the succeeding frame. Thus, the operation is identical to a bank of N x N serial DPCM reconstructors operating in parallel. Both low-frequency subsamples and the high-frequency differences are reconstructed in this fashion.

The final reconstruction step is to generate a low-frequency image version from the subsamples, again using a bilinear interpolation. Finally, the high-frequencies are reinserted into the low-frequency image through the final summation step.
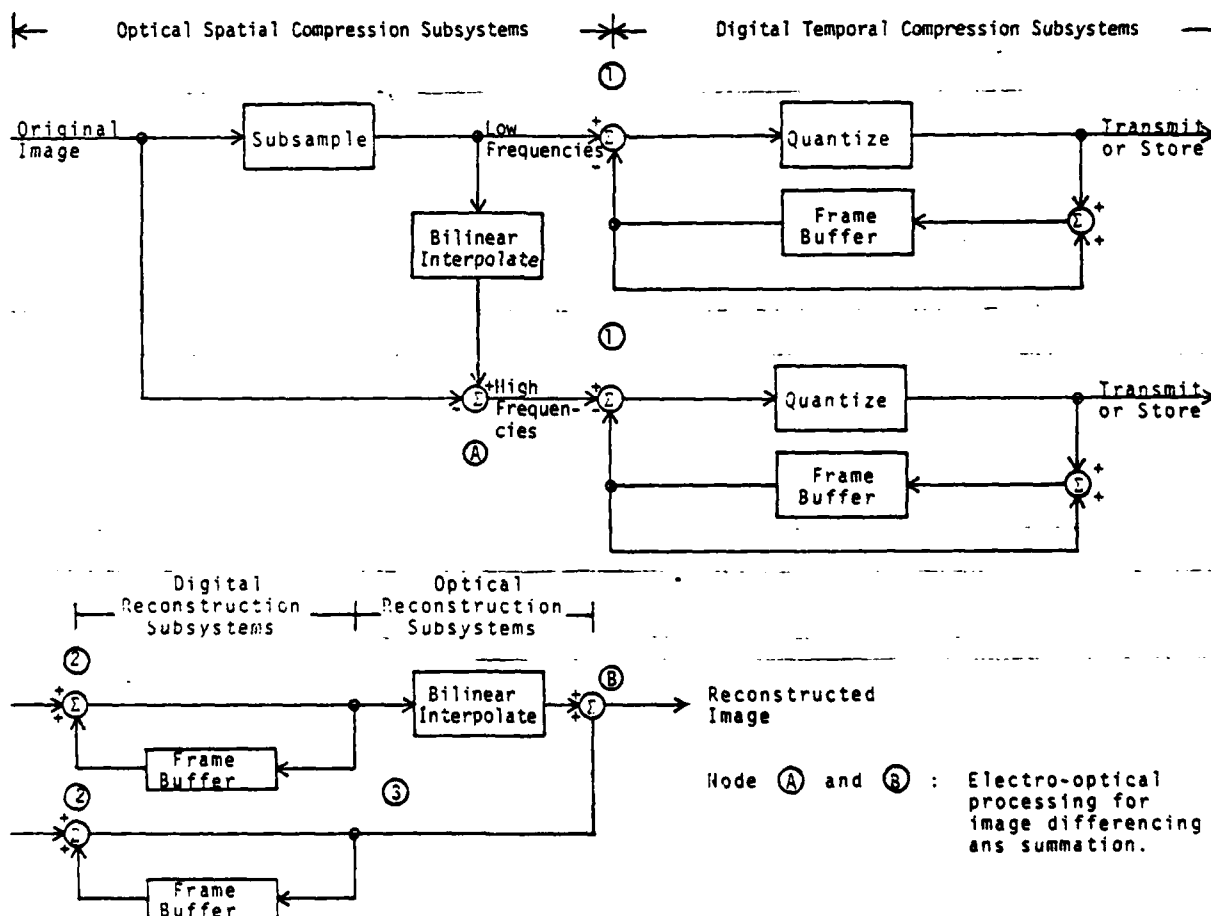

Figure 1: Hybrid (Optical/Digital) Interframe Compression Schematic

Componentry Considerations

The implementation of Figure 1 can be segregated into optical and digital components, with clearly defined interfaces between each. For example, in Figure 1 the nodes marked ① represent where optical information enters a region of the system that is dominated by digital processing. Likewise, in the reconstructor, everything to the right of the point marked ③ is optical processing, with digital processing to the left of ③.

We now summarize the actual component prospects at each of the individual blocks in the diagram.

(1) <u>Subsample</u>. A number of mechanisms can optically subsample an image plane. For example, if a linear sensor array is used, with the array optically "push-broomed" across the image plane, then proper timing of the array read-out can subsample the image plane, e.g., extract every n^{th} pixel from every n^{th} line.

(2) <u>Bilinear interpolation</u>. An optical bilinear interpolation can be constructed by writing the proper apodization function across a lens, and then throwing the lens out of focus [4]. For Figure 1, the subsamples extracted would be imaged on a matrix display, such as a CRT, and the optical interpolation executed from mis-focus on the display.

(3) <u>Image difference</u>. The difference between the original image and the low-frequency version is a pixel-by-pixel difference between the two image planes. Electro-optical mechanisms for this differencing operation have been demonstrated recently, including electro-optical effects in a liquid crystal [5], and the use of channel-plate image intensifiers [6].

(4) <u>DPCM processors</u>. Although the conceptual processes in these components are parallel image-plane to image-plane operations, it is direct to see how they could be implemented serially. For example, subsamples extracted from an array would go directly into the DPCM loop, the loop being a "pass-through" around the quantizer to the frame buffer until the frame buffer is loaded with its reference image. Then the frame buffer data would be extracted, passed into a working quantization loop, etc. The data could be processed serially through a single DPCM processor, the data being serially extracted from the frame buffer and synchronized pixel-for-pixel with the image plane subsamples. Thus, although processed serially through a single DPCM processor, the pixel-by-pixel synchronization between subsamples and frame buffer would have the same effect as N x N parallel DPCM processors operating at a very low data rate. Note that sensing for subsamples with a suitable detector (such as a linear array) could provide data in exactly a suitable format for input to the DPCM processes. Obviously, all operations in the DPCM processors would be digital. Similar comments can be directly applied to the DPCM loop which temporally processes the high-frequencies.

(5) <u>Reconstruction</u>. At the nodes marked ② the incoming data circulates through frame buffers and is summed, with incoming image plane pixels synchronized to the corresponding frame buffer pixels, to regenerate both low and high frequencies.

(6) <u>Reconstruction bilinear interpolator</u>. The processing here is the same as in the compression step. The reconstructed samples would be written on a matrix display which would be imaged out of focus.

(7) <u>Summation</u>. Again, a suitable electro-optical effect would be used to achieve summation, as in the corresponding difference operation in the compression portion of the system. This would require conversion of the samples from the high-frequency DPCM reconstructor into light intensitites for the electro-optical summation.

As can be discerned from this discussion, the overall architure mixes optical and digital processes in a hybrid system for interframe compression.

Simulation Results

To demonstrate the architecture's feasibility, a series of digital simulation experiments was carried out. That is, the optical interpolations were replaced by digital interpolations, the optical differences by digital differences, etc. The resulting digital simulation was carried out in the Digital Image Analysis Laboratory of the University of Arizona, Department of Systems Engineering.

The source data for the simulation consisted of a sequence of 14 digitized frames from a television broadcast of Walter Cronkite. The frames were digitized at 256 x 256 pixels resolution, with 8 bits of intensity per pixel.

The 14th frame NMSE/BR performances for Walter Cronkite Images are summarized in Table 1, and the NMSE performances across each frame for various bit rates are given by Graph 1.1 to 1.4, where the normalized MSE(NMSE) as an objective image quality measure and the overall image quality measure and overall bit rate (BR) as system performace measure are defined as follows:

$$(1) \quad NMSE = \frac{\sum\limits_{x,y} \left[f_r(x,y) - f(x,y) \right]^2}{\sum\limits_{x,y} f(x,y)^2}$$

$$(2) \quad BR = \frac{1}{4} \log_2(LFQ) + \frac{3}{4} \log_2(HFQ)$$

LFQ: low frequency quantization level

HFQ: high frequency quantization level

According to Table 1, the NMSE performances are almost parallel to the LFQ levels. In other words, for each LFQ level, NMSE performances are almost the same within the range of 1% or less. Also, the HFQ levels have little effect on the NMSE performances of the 14th frame at the same LFQ level. On the contrary, for the same HFQ level, the NMSE performances improve considerably along with the increasing LFQ levels. From the transitions of NMSE per-performances across each frame for various combinations of HFQ and LFQ as shown in Graph 1.1 ~ 1.4, it can be said that it is essential to allow the large LFQ levels (8 or 16) for the acceptable objective image quality, regardless of the HFQ levels. The subjective image qualities associated with 1.5 bits/pixel (HFQ = 2, LFQ = 8) and 1.75 bits/pixel (HFQ = 2, LFQ = 16) are quite good as shown in Figures 2 and 3. Figures 4 and 5 show the 14th frame reconstructions for the cases of BR = 1.25 bits/pixel (HFQ = 2, LFQ = 4) and BR = 1.0 bits/pixel (HFQ = 2, LFQ = 2), in which some artifacts due to the motion displacements are observed around his shoulder and his head, and at the center of his face and his chin.

In addition, the transitions of NMSE performances for BR = 1.0 bit/pixel in Graph 1.1 shows very good NMSE performances about 1 ~ 2% up to the 8th frame. Due to the large motion involved between the 8th and 9th frames, the subsequent NMSE performances are deteriorated rapidly. Thus, if we allow the larger LFQ levels such as 8 or 16 at the 9th frame, the errors due to the motion displacement can be made less than the case of LFQ = 2. In other words, a temporally adaptive quantization scheme can improve the NMSE performances further more at the relatively small cost of the BR performances.

Conclusions and Further Research

The conclusions and further research areas can be summarized as follows:

Conclusions

(1) The reasonable subjective and objective image qualities (NMSE - 0.9% ~ 1.3% for the 14th frame of Walter Cronkite images are obtained at the BR = 1.75 bits/ pixel or 1.5 bits/pixel through the digital simulations of the proposed hybrid interframe data compression scheme.

(2) The multi-frame subjective and objective image qualities are mainly determined by the magnitude of the motion displacement between frames. In the proposed hybrid interframe system, it is essential to allow the large LFQ levels (8 or 16) for the acceptable image quality, regardless of the HFQ levels, because the large motion is constituted in the low frequency component. It seems that the contribution of the high frequency component to the reconstructed frame is jeopardized by the motion displacement between frames.

Future research

(1) Development of a Motion displacement measure and motion detector.

(2) Temporally adaptive scheme based on the motion displacement measure.

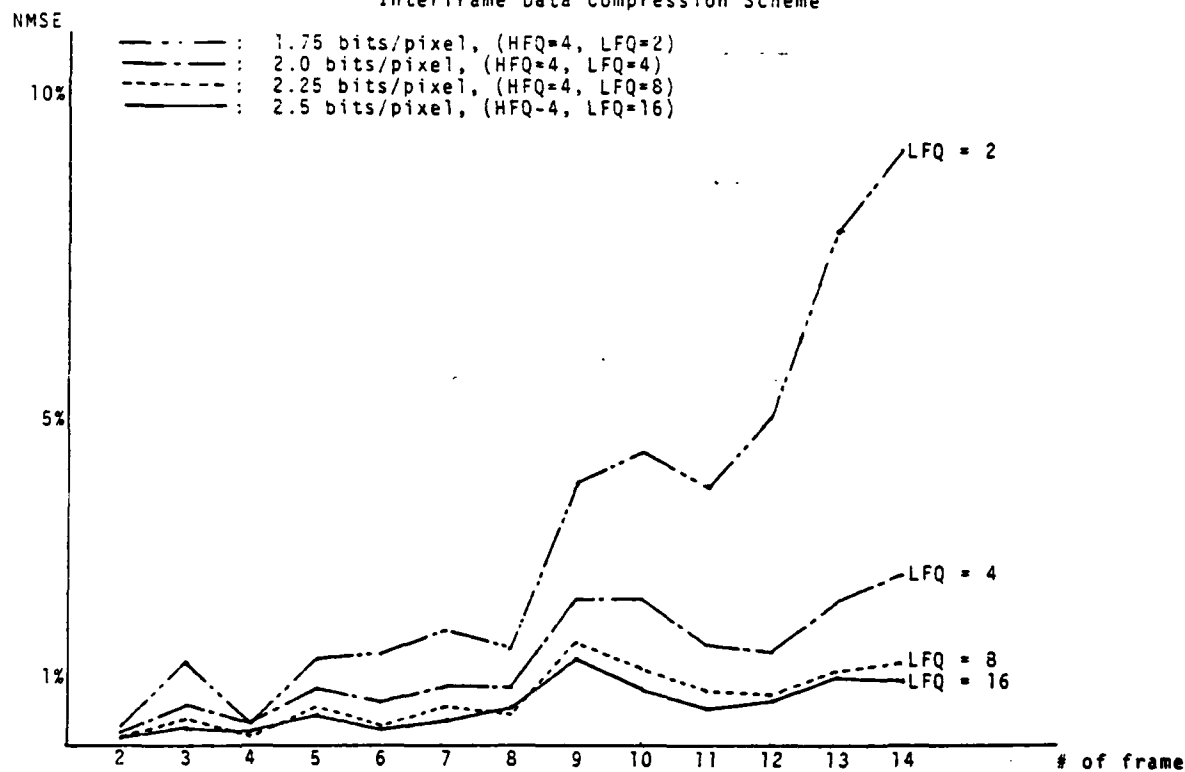(3) Detailed system componentry consideration: Electrical-Optical devices.

Table 1. 14th Frame NMSE/BR Performances
Summary for Walter Cronkite Images

HFQ LFQ	16	8	4	2
16	BR=4.0 bits/pixel NMSE=0.89%	BR=3.25 bits/pixel NMSE=0.83%	BR=2.5 bits/pixel NMSE=0.96%	BR=1.75 bits/pixel NMSE=0.92%
8	BR=3.75 bits/pixel NMSE=1.22%	BR=3.0 bits/pixel NMSE=1.17%	BR=2.25 bits/pixel NMSE=1.24%	BR=1.5 bits/pixel NMSE=1.3%
4	BR=3.5 bits/pixel NMSE=2.77%	BR=2.75 bits/pixel NMSE=2.89%	BR=2.0 bits/pixel NMSE=2.6%	BP=125 bits/pixel NMSE=3.01%
2	BR=3.25 bits/pixel NMSE=9.63%	BR=2.50 bits/pixel NMSE=10.07%	BR=1.75 bits/pixel NMSE=9.15%	BR=1 bit/pixel NMSE=9.53%

Simulation Results for Hybrid (Optical/Digital
Interframe Data Compression Scheme
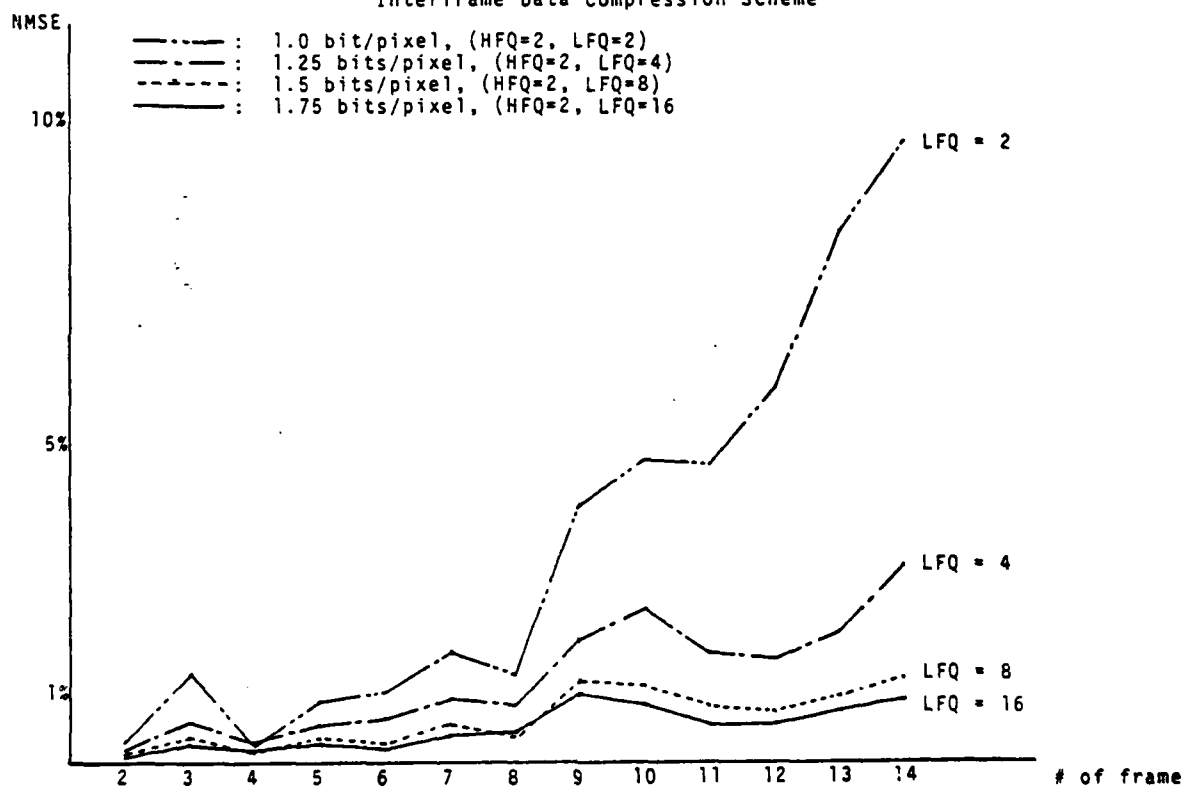
Graph 1.1 Comparison of NMSE Performances for HFQ = 2

Simulation Results for Hybrid (Optical/Digital)
Interframe Data Compression Scheme

NMSE

—·—: 1.75 bits/pixel, (HFQ=4, LFQ=2)
———: 2.0 bits/pixel, (HFQ=4, LFQ=4)
-----: 2.25 bits/pixel, (HFQ=4, LFQ=8)
———: 2.5 bits/pixel, (HFQ-4, LFQ=16)

10%

5%

1%

LFQ = 2

LFQ = 4

LFQ = 8
LFQ = 16

2 3 4 5 6 7 8 9 10 11 12 13 14   # of frame

Graph 1.2 Comparison of NMSE Performances for HFQ = 4

Figure 2. HFQ=2, LFQ=16
BR=1.75 bits/pixel

Figure 3. HFQ=2, LFQ=8
BR=1.5 bits/pixel

Simulation Results for Hybrid (Optical/Digital)
Interframe Data Compression Scheme

NMSE

——·——: 2.5 bits/pixel, (HFQ=8, LFQ=2)
———·———: 2.75 bits/pixel, (HFQ=8, LFQ=4)
---------: 3.0 bits/pixel, (HFQ=8, LFQ=8)
—————: 3.25 bits/pixel, (HFQ=8, LFQ=16)

LFQ = 2

10%

5%

LFQ = 4

LFQ = 8
LFQ = 16

1%

2   3   4   5   6   7   8   9   10 11 12 13 14   # of frame

Graph 1.3 Comparison of NMSE Performances for HFQ = 8
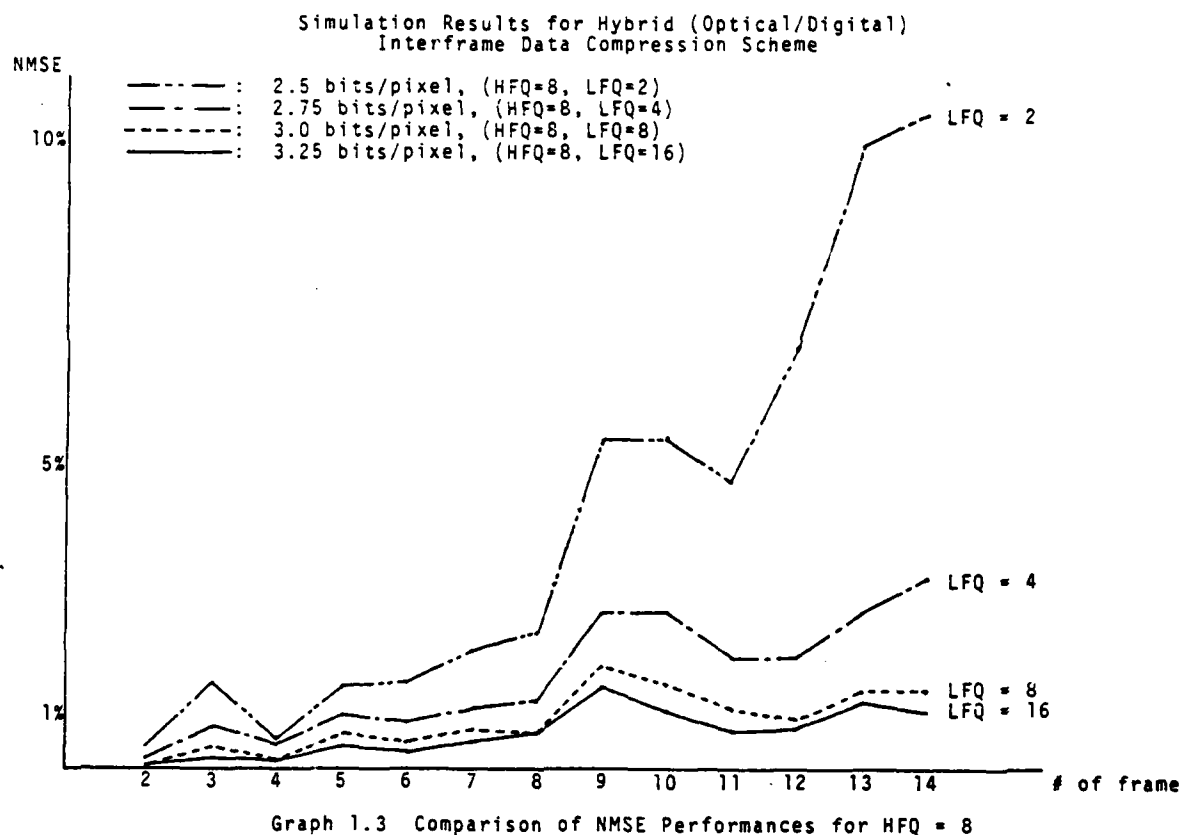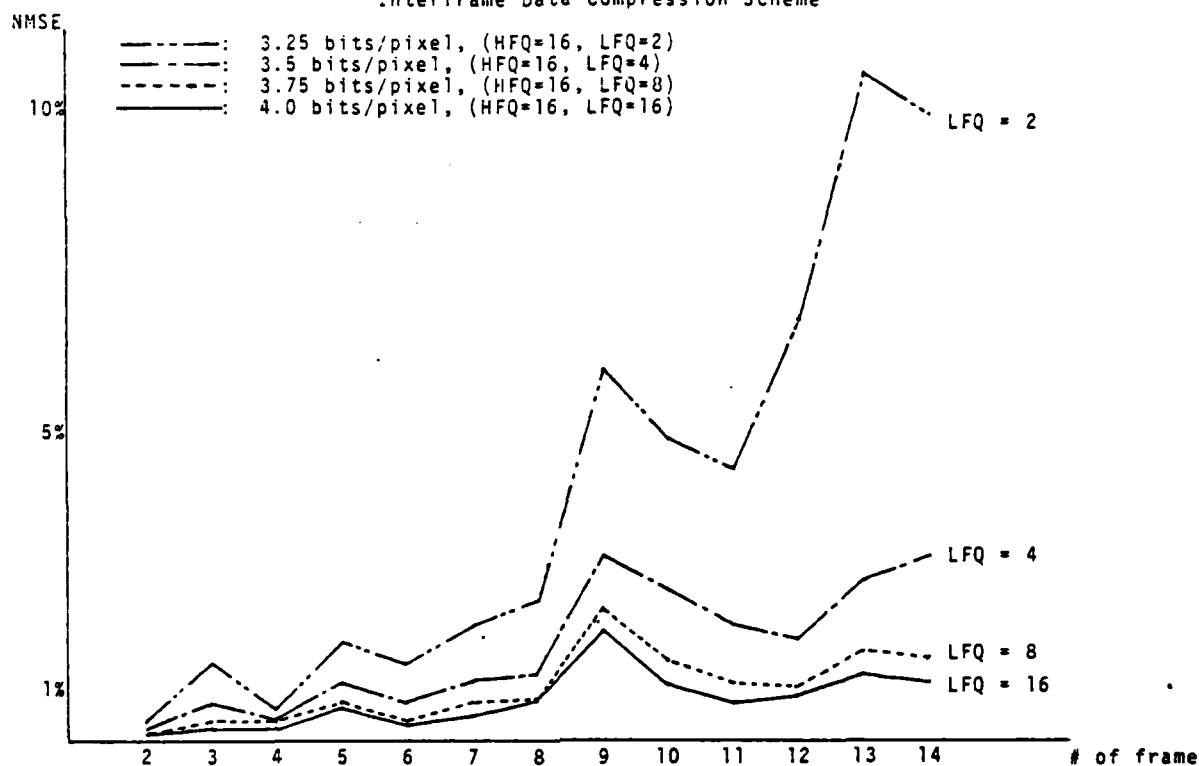

Figure 4.   HFQ=2, LFQ=4
            BR=1.25 bits/pixel

Figure 5.   HFQ=2, LFQ=2
            BR=1.0 bits/pixel

Simulation Results for Hybrid (Optical/Digital)
Interframe Data Compression Scheme

Graph 1.4 Comparison of NMSE Performances for HFQ = 16

Acknowledgements

References

1. B. R. Hunt, "Optical Computing for Image Bandwidth Compression: Analysis and Simulation", Applied Optics, Vol. 17, pp.2944-2951, 1978.
2. W. K. Pratt, Digital Image Processing, Wiley, New York, 1978.
3. B. R. Hunt, "Digital Image Processing", Chapter 4 of Applications of Digital Signal Processing, A. V. Oppenheim, ed., Prentice-Hall, Englewood-Cliffs, New Jersey, 1978.
4. J. W. Goodman, Introduction to Fourier Optics, McGraw-Hill, New York, 1968.
5. J. Grinberg, E. Marom, "Optical Subtraction of Images in Real Time", SPIE Proceedings, Vol. 118, pp. 66-74, 1977.
6. A. Seko, H. Kobayashi, K. Shimizu, "A High-Speed Image Information Preprocessor Using a Channel Plate Image Intensifier", Optical Engineering, Vol. 19, pp. 229-232, 1980.